



A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
 A:Reference number: AB1000; MUID:2017575  
 A:Accession: B81040  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-298 <TEXT>  
 A:Cross-references: GB:AE002530; GB:AE002098; NID:g7227054; PIDN:AAF42138.1; PID:g7227054  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1801

alignment\_scores:  
 Quality: 1469.50 Length: 298  
 Ratio: 5.085 Gaps: 1  
 Percent Similarity: 96.980 Percent Identity: 94.295

alignment\_block:  
 us-09-303-518D-571 x B81040 ..

Align seg 1/1 to: B81040 from: 1 to: 298

```

1 ATGTTTCGTTTACATTCAGAGCTGTTCCCTTTGGCAACCGCATGCA 50
|||||
1 MetPheArgLeuGlnPheArgLeuPheProPheLeuArgThrAlaMetChi 17
51 CATCCGTGTGACCGCCCTGCTCAATGACCTCCCTGCTGCTGCTTCT 100
|||||
17 sIleuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuProLeuSerC 34
101 GTCTGCACAGCTGGGAAACCGCTGGACATCTGGCGTTTACCTTTA 150
|||||
34 yslLeuHisThrLeuGlnLysAsnArgLeuGlnLysLeuAlaPheTyrLeuLeu 50
151 AAGGAAGACCGCGCGCATCTGCGCAATATGCGCGAGCGGGTTGAA 200
|||||
51 LysGlnAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyLeuAs 67
201 CCCGACACGACGAGGTCAAGCGCTTTTGGGAAACCGCAAAATGGG 250
|||||
67 nProAspProLysThrValLysAlaValPheAlaGlnThrAlaLysGly 84
251 GTTGGAGACTTGGCCCGCGCTTTTCAAAAACCGGAAGCATCGAAAC 300
|||||
84 LysLeuGlnLeuAlaPheAlaPheArgLeuPheProLeuArgIleGlnThr 100
301 ATGTTCAAGCGGTACACGCTGGGAAACGCTGACAGCGCTTTGGCAA 350
|||||
101 MetPheLysAlaValHisGlyTyrGlnHisValGlnGlnAlaLeuAspLys 117
351 GGGCGAAGGCGTGTTCATCAGCGCGCATCGGACATCGGAGTACGATT 400
|||||
117 shIsugLysLeuLeuPheIleThrProHisIleGlySerTyrAspLeuG 134
401 GCGGAGCTACATCAGCAGCAGCTTCGCTTCACCTGACCGCATGAC 450
|||||
134 LysIleArgTyrIleSerGlnGlnLeuProPheProLeuThrAlaMetTyr 150
451 AAGCGCGCAAAATCAAGCAGTACGAAATCATGACGCGCGGAGGCT 500
|||||
151 LysProPheLysIleLysAlaIleAspLysIleMetGlnAlaGlyArgVa 167
167 LArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGln 184
551 TCATCAAGCGCTTCGCGCGCGGAGGACCATCATCTGCGCCACACAC 600
|||||
184 IleIleLysAlaLeuArgSerGlyGlnAlaThrIleValLeuProAspHis 200
601 GTCCCTTCCTCCGAGAGGCGCG...GCGCTGTGGCGGATTTTTCGG 647
|||||
201 ValProSerProGlnGlnGlyLysGlnGlyValTyrPheValAspPhePhe 217

```

```

648 CAACCTTCATACACATGACCTGGCGGCAAAATTTGSCACAGCTCAAG 697
|||||
217 yLysProAlaTyrThrMetThrLeuAlaAlaLysLeuAlaHisValLysG 234
698 GCGTGAAGAACCTGTTTCTGCTGCGACGCTGCCCGACGACAAAGCG 747
|||||
234 LysAlaLysThrLeuPhePheCysGlnLysLeuProGlnGlyGlnGly 250
748 TTGCTGTGTGACATCCCGCGCTTCGCAAGGGAATTTGACGCAACAAG 797
|||||
251 PheAspLeuHisIleArgProValGlnGlyLysLeuAsnGlyAspLysAl 267
798 CCACGATCGCGCGCTGTTCACACCGCAATACCGAATATGATACCGCG 847
|||||
267 AhIsAspAlaAlaValPheAsnArgAsnAlaGlnTyrTyrPheArgTyr 284
848 TTTCGACGCGATGCTGTTTATGTCACACCGCTTAAACGCG 891
|||||
284 heProThrGlnTyrLeuPheMetTyrAsnArgTyrLysMetPro 298

```

seq\_name: p1r2:D81986

seq\_documentation\_block:  
 probable acyltransferase NMA0662 [imported] - *Neisseria meningitidis* (strain Z2491 se  
 C:Species: *Neisseria meningitidis*  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #extl\_change 02-Feb-2001  
 R:Accession: D81986  
 R:Parthill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo  
 R:Holroyd, S.; Jajels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre  
 A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491  
 A:Reference number: AB1775; MUID:2022556  
 A:Accession: D81986  
 A:Molecule type: DNA  
 A:Status: preliminary  
 A:Residues: 1-298 <PAR>  
 A:Cross-references: GB:AL62753; GB:AL157959; NID:g7379120; PIDN:CA83949.1; PID:g737  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: NMA0662

alignment\_scores:  
 Quality: 1467.50 Length: 298  
 Ratio: 5.078 Gaps: 1  
 Percent Similarity: 96.980 Percent Identity: 93.960

alignment\_block:  
 us-09-303-518D-571 x D81986 ..

Align seg 1/1 to: D81986 from: 1 to: 298

```

1 ATGTTTCGTTTACATTCAGAGCTGTTCCCTTTGGCAACCGCATGCA 50
|||||
1 MetPheArgLeuGlnPheArgLeuPheProPheLeuArgThrAlaMetChi 17
51 CATCCGTGTGACCGCCCTGCTCAATGACCTCCCTGCTGCTGCTTCT 100
|||||
17 sIleuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuProLeuSerC 34
101 GTCTGCACAGCTGGGAAACCGCTGGACATCTGGCGTTTACCTTTA 150
|||||
101 MetPheLysAlaValHisGlyTyrGlnHisValGlnGlnAlaLeuAspLys 150
34 yslLeuHisThrLeuGlnLysAsnArgLeuGlnLysLeuAlaPheTyrLeuLeu 50
151 AAGGAAGACCGCGCGCATCTGCGCAATATGCGGACGCGGTTTGA 200
|||||
51 LysGlnAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyMetAs 67
201 CCCGACACGACGAGGTCAAGCGCTTTTGGGAAACCGCAAAATGCG 250
|||||
67 nProAspProLysThrValLysAlaValPheAlaGlnThrAlaLysGly 84

```

```

251 GTTGGAACTGCCCCCGCTTTTTCAAAAACGGAGACATCGAACA 300
|||||
84 LYLGGILLeuAlaProAlaPhePheArgLysProGlnAspIleGluThr 100
|||||
301 ATGTTCAAAGCGGTACACGGCTGGGAACAGCTGACAGCGTTTGACAA 350
|||||
101 MetPheLysAlaValHisGlyTrpGlnHisValGlnGlnAlaLeuAspLys 117
|||||
351 GGGGAAAGGCGTGTCTTCATCAGCGCCGACATCGGACACTTCGATTGG 400
|||||
117 SHLSGLGLLeuLeuPheLeuPheLeuPheHisIleGlySerLysAspLeuG 134
|||||
401 GGGAGCGTACATCAGCCAGACAGCTTCGTTCCACTGACCGCATGTAC 450
|||||
134 LYLALGlyLysIleSerGlnGlnLeuProPheProLeuThrAlaMetLys 150
|||||
451 AAGCGCCGAAATCAAGCATAGACAAATCATGCAAGGGGCGAGGCT 500
|||||
151 LysProProLysIleLysAlaIleAspLysIleMetGlnAlaGlyArgVa 167
|||||
501 GGGCGCAAGCGCAAAACCGCGCCGACGATACAAAGGGGTCAACAA 550
|||||
167 LAGGLYLSGLYLSThrAlaProThrSerIleGlnGlyValLysGlnI 184
|||||
551 TCATCAAGGCGCTGGCGCGCGGAGGAGGACCATCATCTCCCGGACAC 600
|||||
184 LeIleLysAlaLeuArgSerGlyGlnAlaThrIleValLeuProAspHis 200
|||||
601 GTCCCTTCCTCCGAGAAAGCGGC...GCGCTGTGGCGGATTTTTCGG 647
|||||
201 ValProSerProGlnGlnGlyGlyGlyValThrValAspPhePheGln 217
|||||
648 CAACCTGCATACACATGACATGCGGCGCAAAATTTGGCACAGCTCAAG 697
|||||
217 LYLSPProAlaLysThrMetThrLeuAlaAlaLysLeuAlaHisValLysG 234
|||||
698 GCGTGAAGACCGCTTTTGTCTGCGCAAGCGCTCCGACGAGCAAGG 747
|||||
234 LYLALYLSThrLeuPhePheCysGlyGlnArgLeuProGlyGlnGly 250
|||||
748 TTCGTGTGGACATCCGCGCGTCGCAAGGGAATTTACAGCGCAACAA 797
|||||
251 PheAspLeuHisIleArgProValGlnGlyGlnLeuAsnGlyAspLysAl 267
|||||
798 CCAGCATGCCCGCGCTGTCAACCGCAATACCGAATTTGATAGCGCGT 847
|||||
267 AHLSAspAlaAlaValAlaPheAsnArgAsnAlaGluTrpIleAlaArgP 284
|||||
848 TTCGACCGCGATCTGTTATGTACACCGCTATPAAACGCGC 891
|||||
284 heProThrGlnTrpLeuPheMetLysAsnArgLysLysMetPro 298

```

```

seq_name: p1r2:H83642
seq_documentation_block:
  Probable 2-OH-lauricoyltransferase PA0011 [Imported] - Pseudomonas aeruginosa (strain PA01
  C:Species: Pseudomonas aeruginosa
  C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
  C:Accession: H83642
  R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
  admin, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
  Nature 406, 959-964, 2000
  A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
  A:Reference number: A82950; M01D:20437337
  A:Accession: H83642
  A:Status: Preliminary
  A:Molecule type: DNA
  A:Residues: 1-295 <STO>
  A:Cross-references: GB:AE004441; GB:AE004091; NID:99945828; PIDN:AA03401.1; GSPDB:GN001
  A:Experimental source: strain PA01
  C:Genetics:
  A:Gene: PA0011

```

## alignment\_scores:

Quality:	266.50	Length:	285
Ratio:	1.514	Gaps:	7
Percent Similarity:	61.754	Percent Identity:	26.667

## alignment\_block:

US-09-303-518d-571 x H83642

Align seg 1/1 to: H83642 from: 1 to: 295

```

55 CTGTGACCGCCCTGCTCAATGCTTCCTTCCTGCTGCTTCCTGCTCT 104
|||||
8 LeuValValGlyAlaLeuArgLeuPheAlaLeuLeuProTrpArgAlaVal 24
|||||
105 GCACACGCTGGGAAACCGGCTCGACATGTGGCGTTCCTTAAGG 154
|||||
24 LGLNGlyValGlyAlaGlyIleGlyTrpLeuMetTrpLysLeu.....P 39
|||||
155 AAGACCGCGCGCGCATGCGGCAATATGCGGAGCGGCTTGAACCC 204
|||||
39 roAsnArgSerArgGlnValValArgIleAsnLeuSerLysCysPhePro 55
|||||
205 .....GACACGACGCGTCAAGCGCTTTTCGCGAAACCGCAA 245
|||||
56 GluLeuSerGlnThrGlnLeuGlnLysLeuValGlnIleSerLeuMetAs 72
|||||
246 ATGCGGT.....TTGCACTTGCCCGCGCTTTTCAAAACCGG 286
|||||
72 PLeuGlyArgThrLeuThrGlnSerAlaCysAlaTrpIleTrpProGln 89
|||||
287 AAGCATGCAAAATGTTCAAGCGGTACACGCTGGGCAACGCGGAG 336
|||||
89 LulysSerLeuArgTrpIleArgGlnValGlnIleMetGlnValLeuGln 105
|||||
337 CAGCCTTTGGCAAGGCGGCAAGGCTGCTGTATACGCGGCACATCGG 386
|||||
106 GlnAlaLeuAlaSerGlyAspGlyLeuValGlyIleThrSerHisLeuGln 122
|||||
387 CAGCTACGAT...TTGGCGGACGTACATCAGCGCGACGTTCCGTTCC 433
|||||
122 YAsnTrpGlnValLeuAsnHisPheTrpLysSerTrpAlaLysProIle 138
|||||
434 ACCTGACCGCGATGTCAAGCGCGCGCAAAATGCAACGATACAAATC 483
|||||
139 .....LLeuPheLysArgProProLysLeuLysAlaValAspLysLeu 152
|||||
484 ATGACAGCGCGGAGCGGTGCGCGCAAAAGGCAAAACCGCGCCACGCGAT 533
|||||
153 LeuLysLysGlnArgValGlnLeuGlnLysAsnArgValAlaProSerThrPr 169
|||||
534 ACAAGGGGTCAAAACAATCATCAAGGCGCTGCGCGCGGCGGCAACGA 583
|||||
169 OGluGlyIleLeuSerValIleLysGlnValLysLysLysLysValG 186
|||||
584 TCATCTCGCCGACGACGCTCCCTTCGCGAGAGCGGCGGCTGTGG 633
|||||
186 LYLLeuProAlaAsp.....ProGlnProAlaArgTrpAlaGlyLeuPhe 200
|||||
634 GCGGATTTTTCGCAAACTGTCATACACCATGACCTGGCGGCAAAAT 683
|||||
201 ValProLysLeuGlyThrThrAlaLeuIleSerLysPheValProGlnLe 217
|||||
684 GGCACACGTCAAAAGCGGCAAAACCTGTTTCTGCTGCGAAGCGCTGC 733
|||||
217 LLeuSerArgGlyLysAlaArgGlyValAlaPhePheHisAlaValArgLeuP 234
|||||
734 CCGAGGACAAGCGCTGCTGTGACATCCGCGCGCGCAAGGGGATTTG 783
|||||
234 roAspGlySerGlyTrpLysValIleLeuGlnAlaAlaProAlaAspMet 250
|||||
784 AAGCGCAACAAAGCGCAC...GATGCCCGCGGTGTTCAACGCAATACGGA 830

```

```

251 TyrAspLysAspLeuAlaValSerValAlaAlaMetSerArgGluLeuAl 267
      ::::: |||:||||:||||:||||:
831 ATATTGGATACCGCGTTTCGACGACGATTCGTATATGATACACCGCT 880
      ::::: |||:||||:||||:||||:||||:
267 alystValAlaValAlaValProSerGlnTyrMetTrpSerMetLysArg 284
      ::::: |||:||||:||||:||||:
881 ATAAA 885
      ::::: |||:
284 helys 285

```

seq\_name: p1r2:C82691

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase XP1348 [imported] - Xylella fastidiosa (str C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: C82691  
R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen-  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: C82691  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-316 <STM>  
A:Cross-references: GB:AE003967; GB:AE003849; NID:99106347; PIDN:AAE84157.1; GSPDB:GN001  
A:Experimental source: strain 945C  
R:Simpson, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A-  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H-  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincan, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm-  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klieger, J.E.; Krumme, E.E.; Laly-  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E-  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XP1348

alignment\_scores:

Quality:	258.00	Length:	283
Ratio:	1.518	Gaps:	9
Percent Similarity:	60.071	Percent Identity:	27.915

alignment\_block:

US-09-303-518d-571 x C82691 ..

```

Align seg 1/1 to: C82691 from: 1 to: 316
103 CTGACACAGCGTGGGAACCGCTCGACATCTGCGTTTACCTTTAA 152
      |||:||||:||||:||||:||||:
10 LeuTyrArgValAlaSerLeuValGlyHisIleProTyrProPheLeu 26
      ::::: |||:||||:||||:||||:
153 GGAAGACCGCGCGCATCGTCCG.....AATATGC 184
      ::::: |||:||||:||||:||||:
26 sArg...LeuSerAspLeuLeuAlaTyrCysTrpLeuLysLeuAsnAla 42
      ::::: |||:||||:||||:||||:
185 GGCAGCG.....GGTTG 198
      ::::: |||:||||:||||:||||:
42 rglGlsrArgValThrGlnArgAsnLeuGluLeuThrTyrProGluLeu 58
      ::::: |||:||||:||||:||||:
199 AACCCCGAC.....ACGACAGCGTCAAGCCGTTTTCGCGGAACGCG 242
      ::::: |||:||||:||||:||||:
59 SerProGlnGlnHisThrArgLeuHisGlnIleIleTyrSerThrVal 75

```

```

243 AAAATGGCGTTTGACCTTCCCGCGGTTTTCAGAAAAACCGAAGACA 292
      ::::: |||:||||:||||:||||:
75 lArgGlnAlaPheGluMetLeuHisIleThrPheHisProGluThrGlu 92
      ::::: |||:||||:||||:||||:
293 TCGAACAATGTTCAAAAGCGGTACACCGCTGGAGACACGTGCACAGGT 342
      ::::: |||:||||:||||:||||:
92 snLeuAlaValLeuArgGluTyrHisGlyValAspLeuTyrSerAlaVal 108
      ::::: |||:||||:||||:||||:
343 TTGACAAAGCGCGAAGGCTCTGTTTCATCAGCGCGACATCGGACGTA 392
      ::::: |||:||||:||||:||||:
109 LeuAlaArgAsnArgGlyMetIleIleValAlaProHisPheLysAsn 125
      ::::: |||:||||:||||:||||:
393 CGATTGGCGCGGCTATACATCACCGACGACGCTTCCTACCTGACCG 442
      ::::: |||:||||:||||:||||:
125 polLeuLeuLysGlnTyrLeuSerAspLeuAlaPro.....LeuThr 140
      ::::: |||:||||:||||:||||:
443 CCATGTACAGCGCGCGAATAATCAAGCGATAGCAAAATCATGACGCG 492
      ::::: |||:||||:||||:||||:
140 eumTyrArgProAlaAsnSerAlaValLeuAspGlyPheLeuGln... 155
      ::::: |||:||||:||||:||||:
493 GGCAGGCGCGCGCAAGCAAGCAAGCGCGCGCGGCGGACATCAAGG.. 540
      ::::: |||:||||:||||:||||:
156 ...ArgValArgGlyGlyAsnAsnValHisGlnValAlaGlnGluGly 171
      ::::: |||:||||:||||:||||:
541 ...GTCAACAATCATCAAGCGCGCTTCGCGCGCGCGCAATTCG 586
      ::::: |||:||||:||||:||||:
171 talValArgHisLeuPheLysValLeuLysAsnGlyThrIleGly 188
      ::::: |||:||||:||||:||||:
587 TCCTGGCGCGCACGCGCTTCGCGCGAGAGCGCGCGGTGGCG 636
      ::::: |||:||||:||||:||||:
188 leuLeuProAspGln.....GlnProLysIleGlyGlnGlyValPheAla 202
      ::::: |||:||||:||||:||||:
637 GATTGTTGGCGCAACCTGATACACCGATACGACGCGCGCGCAATTCG 686
      ::::: |||:||||:||||:||||:
203 ProPhePheGlyIleGluAlaLeuThrMetThrLeuValGlnArgLeuAl 219
      ::::: |||:||||:||||:||||:
687 ACAGCTCAAGCGCGTGAACACCTGTTTTCGCTCCGACCGCTCGCG 736
      ::::: |||:||||:||||:||||:
219 aGluArgThrSerAlaThrValLeuTyrAlaTrpCysGluArgIleSerP 236
      ::::: |||:||||:||||:||||:
737 ACGGACAGCGCTTCGTTTCACATCCGCGCGCGCGCGCAAGGAAATG... 783
      ::::: |||:||||:||||:||||:
236 rGlnIleGluPheAlaLeuHisMetGlnHisAlaAspProAlaValAla 252
      ::::: |||:||||:||||:||||:
784 AACGGCAACAAAGCCGACGATGCCCGCGTTCACCGCATTCGAAATA 833
      ::::: |||:||||:||||:||||:
253 AsProAspProLeuIleAlaIleAlaThrAlaLeuAsnAlaGlyIleGlu 269
      ::::: |||:||||:||||:||||:
834 TTGATACCGCGTTTCGCGACGATCTGTTTATGATACACCGCTAT 882
      ::::: |||:||||:||||:||||:
269 glleAlaArgArgAspProThrGlnTyrGlnThrTyrLysArgTyr 285
      ::::: |||:||||:||||:||||:

```

seq\_name: p1r2:D83239

seq\_documentation\_block:

probable lauroyl acyltransferase PA3242 [imported] - Pseudomonas aeruginosa (strain P  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83239  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; L-  
adman, S.; Yun, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lathig, K.; L-  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa-  
A:Reference number: A82950; MUID:20437337  
A:Accession: D83239  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-312 <STM>  
A:Cross-references: GB:AE004747; GB:AE004091; NID:9949362; PIDN:ANG06630.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3242

```

72 CAATGCGCTCCCTGCTGTGCGCTTTCGTGTGCACACGGCGGAACG 121
31 GlnIleuProIytrProValIleuAargPheIeuGlyIthrArgYthnGlyIysLe 47
122 GCGTCGGACACATCTGGCGTTCCTTACCTTTTAAGGAGA
47 uAlaIarg.....ProheIeuLysrArgIgluSerIlea 59
159 .....CCGCGCGCGACAGT 173
59 IagIlnYsaInIIegIulbeCysPheProThrIleuSerIArgIgluIuArg 75
174 CGCCAAATATGCGGACGGCGGTTTGTAACCCGACAGCAGACGGGTCAAG 223
76 GlnIysLeuIleIaIaIuInasphe.....HisserIeuGlymet.A 89
224 CGGTTTTTTCGGGAAACGGAATATGCGTTTGAGAACTGGGCCCCGGCTTT 273
89 ILeu.....LeuGlnIthrIgluMetAlaIarp 97
274 TTCAAAAACCGGAGACACATCGAAACATGTTCACAAACGGGTACACGGCG 323
98 PheItrProaspSerIArgValAlaIArgIytrPheIasp.....ValAspIyIys 113

```

324 GGAACACGTCAGCAGGCTTTGGACAAAGGCGCTGCTGTATCA 373  
 113 uaspsnleuthrargalaglnasnaarglyvalmetvalalg 130  
 374 CGCGGCATCGCAGCTACGATTTGGGCGCTACATCAGCCAGCAG 423  
 130 lvalahisphenetserleugluleuglylarmyalmetylueuys 146  
 424 CTTCGCTTCACCTGACCGCATCTACAGCCGCCAAATCAACGAT 473  
 147 glnpro.....metvalatthrargprohisasnproleuue 161  
 474 AGACAAATCATCGAGCGGCGCGCGCGCAAGGCAAAACCGCGC 523  
 161 tglutrpvalglnhrargglyarmetarg.....serasnysalam 176  
 524 CCACGCGCATACAGGGGTCAAAACATCATCAAGCCCTCGCGCGG 573  
 176 etlleglylarmasnleuarglylvalalglalaleuylslygly 192  
 574 GAGGACCATCATCTCGCGCGCGCGCGCGCGCGCGCGCGCG 623  
 193 glualalvaltrpnhelaprosapelnspyrlyglyprolys.....gl 207  
 624 CGCGCTGTGGCGGATTTTTCGCG.....AACTGCATACACCATGAC 670  
 207 yserserphenalaprohephenalavalgluasnvalalathrthrang 224  
 671 TGCGGCAAAATTTGCAACGCGTCAAGGCGTGAACCTGTTTTCG 720  
 224 lthrtrvalleuuserargleuseglyalalaleuethrvalthr 240  
 721 TGCAAGCGCTGCGCGAGCAAGGCTGCTGCTGACATCCGCCGT 770  
 241 metvalarglysserapsenserglylarmleuthrlethrpro... 256  
 771 CCAAGGGAATTTGACCGCAACAGCCGCGAT.....CGCGCG 811  
 257 .....glmetglulgltylproalaspelnasnglnalalalal 271  
 812 TGTCACCGCATATCCGAATATTGATACCGCTTTTCGAGCGAGTAT 861  
 271 ymetaleuylsillelleuglulmetargalaprogluglntyr 287  
 862 CTGTTATGTACAACCGCTATATAACG 888  
 288 leutrpilenhisargargphenylsthr 296  
 seq\_name: pliz: A10441  
 seq\_documentation\_block:  
 probable membrane protein YP03632 [imported] - Yersinia pestis (strain C092)  
 C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
 C:Accession: A10441  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
 deno-Tarraga, A.M.; Chillingworth, T.; Crokin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: A10441  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-306 <KUP>  
 A:Cross-references: GB:AL590842; PIDN:CAN93101.1; PID:gl5981553; GSPDB:GN00175  
 A:Genetics:  
 A:Gene: YP03632

alignment\_scores: 207.50 Length: 287

Ratio: 1.289 Gaps: 9  
 Percent Similarity: 56.098 Percent Identity: 28.920  
 alignment\_block:  
 US-09-303-518d-571 x A10441 ..  
 Align seg 1/1 to: A10441 from: 1 to: 306

67 CTGCTCAAAATGCCCTCCCTGCTGCTGCTTCTGTCACACGCTGG 116  
 25 Valleuethleuvalalglulnleuprotlyrproleuethleuval 41  
 117 AAACCGGCTCGGACATCGCGCTTTTACCTTTAAAGAAAGCGCGCG 166  
 41 yvaltrpneuglylarmthrsermetargphenyls.....argarg 56  
 167 GCATCGTCGCAATATCGCGAGCGGCTTTGAACCCGAC..... 207  
 56 alserllealargargasnleugluleucyspherprouaspmetaplys 72  
 208 .....ACGACAGCGTCAAAAGCGCTTTTTCGGAACGCGCAATG 248  
 73 glvalleuglulnthrillelleglyasnphelglserleuethleu 89  
 249 CGGTTTGGACCTTGGCGCGCGCTTTTCAAAAACCGGACATCGAA 298  
 89 yleuethleuethrlymetaleatrpheerpservatpargllegln 106  
 299 CATGTTCAAAAGCGGTACACGCGCTGGGACACGTCGACAGCTGAT 348  
 106 rgtprphe.....servalserglyleuethleuvalarglalegln 121  
 349 AAGGCGAAGGCTGCTGCTTTCATCGCGCGCATCGGACATCGAT 398  
 122 glylarmarglylvalleuvalilleglyalhisphenetserleu 138  
 399 GCGCGAGCGCTACATACGCGCGCGCTTCCGCTGACCGCGCGAT 448  
 138 ugllylarmarglaleuethleuethleuethleuethleuethleu 153  
 449 ACAAGCGCGCGCAAAATCAAAAGCGATGACAAATGATGACGCGG 498  
 153 yratrprouhisasnlyvalmetgluleuvalalthrtrpneugly 169  
 499 GTGCGCGGCAAGGCAAAAGCGCGCGCGCGCGCATGCAAGGCGT 548  
 170 metarg.....serasnlysalmetleuaspargargpdeuargl 184  
 549 AATCATCAAGCGCTGCGCGCGCGCGCGCGCGCGCATGATCGCG 598  
 184 ymetvalarglaleuethleuethleuethleuethleuethleu 201  
 599 ACGTCCCTTTCGCGAGGAGCGCGCGCGCGCGCGCGCGCGAT 648  
 201 lnaerlylglprouarg.....glyservalphealaproleupheal 215  
 649 .....AACTGCATACACCATGACGCGCGCGCGCGCGCGCGCG 695  
 216 valglunlinalalalalthrserglylthrhemeleuethleu 222  
 696 AGGCGTGAAGCCGCTTCTGCTGCGCGCGCGCGCGCGCGCGCG 745  
 232 alysproulaleuethleuethleuethleuethleuethleu 249  
 746 GCTGCTGTTGCATCGCG.....CCGTCGAAGG 777  
 249 ltylraspneuleuethleuethleuethleuethleuethleu 265  
 778 GAATTCACGCGCAACAGCGCGCGCGCGCGCGCGCGCGCGCG 827  
 266 glullealalalalalalalalalalalalalalalalalal 277  
 828 CGAATATTGATACGCGCTTTCGCGCGCGCGCGCGCGCGCGCG 877



277 egluLysgluilemetargalaprogluIntrLeuPheuhisArgA 294  
 878 GCTATATAAGC 888  
 294 rphelystr 297

seq\_name: p1r2:D81085

seq\_documentation\_block:

HXB/MSB family protein NMB1418 [imported] - Neisseria meningitidis (strain MC58 serogroup C)  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: D81085  
 R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vignani, A.; Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: AB1000; M01D:20175755  
 A:Accession: D81085  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-289 <RET>  
 A:Cross-references: GB:AE002491; GB:AE002098; NID:g7226655; PIDN:AAF41779.1; PID:g7226655  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1418

alignment\_scores:

Quality	Ratio	Length	Gaps
204.00	1.207	289	12
Percent Similarity: 58.478		Percent Identity: 27.336	

alignment\_block:

US-09-303-518D-571 x D81085 ..  
 Align seg 1/1 to: D81085 from: 1 to: 289

```

67 CTGCTCAATGCTCTCCCTGCTGCTGCTCTCTCTGTCACACGCTGGG 116
7 ValLeuYrValLeuInPheuProPheAlaLeuHisLysIleAl 23
117 AAACCGGCTCGACATCTGCGCTTTTAAAGGACGCGCGC 166
23 aaspleuthrnglyLeuLeuAla...TyrLeuLeuValLysProArgArgA 39
167 GCATC.....GTGCGCAATATGCGGCGACGCGGTTTGAACCC 204
39 rglleLysgluileasInLeuAlaLysCysPheSerGluTrpSerGlu 55
205 GACACGACGCGTC...AAAGCGCTTTTGGGAAACGCAAAATCGG 251
56 LysArgLysThrValLeuLysGlnHisPheLysHisMetAlaLysLeu 72
252 TTGGAACTGCGCCCGGCTTTTCAAAAACGGAAGACATGCAACAA 301
72 tleuLysgluYrGlyLeuYrTrpYrAlaProAlaGlyArgLeuLysSer 89
302 TGTTCAAAGCGGTACACGCTGGGACACGTCGACAGCGCTTTGACAG 351
89 euValArg...TyrArgAsnLysHisLysIleAspAspAlaLeuAla 104
352 GCGGAAGGCGCTGTCATCAGCGCCGACATCGGACGTCAGATTTGGG 401
105 GlyLysValIleIleLeuYrProHisPheThrAlaPheGluMetAl 121
402 CGGACGCTACATCAGCAGCAGCTTCGTTCACTGACCGCATGTACA 451
121 aValTrpAlaLeuAsnGlnAspIlePro.....LeuIleSerMetLys 136

```

```

452 AGCGCCGCAAAATCAAGCGATAGACAAATCATGACGCGCGCAGGTC 501
136 erHisGlnLysAsnLysIleLeuAspGluGlnIleLeuLysGlyArgAsn 152
502 CGCGCGCAAA.....GGCAAAACCGCGCCCGCCAGGCATACA 536
153 ArgTyrHisAsnValPheLeuIleGlyArgThr.....G 164
537 AGGGGCAAAACATCATCAGCGCTGCGCGCGGCGAGCAACATCA 586
164 uLysLeuArgAlaLeuValLysGlnPheArgLysSerSerAlaProPhe 181
587 TC...CTGCGCGACCGACGCTCCCTCTCTCGCGAGGAAGCGG 624
181 euYrLeuProAsp.....GlnAspPheGlyArgAsnAsp 192
625 GCGCGTGGCGCGATTTTTCGCAAACTGCATACACATGACACTGGC 674
193 SerValPheValAspPhePheGlyIleGlnThrAlaThrIleThrGly 209
675 GCGCAAAATGCGCACACGTCGCAAGCGGTGAACCCGCTTTTCTGCTGCG 724
209 uSerArgIleAlaLeuAlaAsnAlaLysValIleProAlaIlePro 226
725 AAGCGCTGCGCGACGACGACAGCGCTGCTGTCACATCGCGCCGTCAA 774
226 AlaArgGlnAlaAspAsnThr...ValThrLeuHisPheYrProAlaTrp 241
775 GGGCAATGAACGCGCAACAAAGCCAC...GATGCGCGCGCTGTCAACG 821
242 LysSerPheProGlyLysAspAlaLysAlaAspAlaGlnArgMetAsn 258
822 CAATACCGAATATGATACGCGCTTTCGACGACGATGCTGTATGT 871
258 gPheIleGluAspArgValArgGlnHisProGluGlnTrpYrPheTrpLeu 275
872 ACAACGCGCTATTAACG 888
275 IsLysArgPheLysThr 280

```

seq\_name: p1r2:B81857

seq\_documentation\_block:

probable acetyltransferase NMA1630 [imported] - Neisseria meningitidis (strain Z2491)  
 C:Species: Neisseria meningitidis  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
 C:Accession: B81857  
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mouton, R.; Holtroyd, S.; Jorgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre  
 Nature 404, 502-506, 2000  
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
 A:Reference number: AB1775; M01D:20222556  
 A:Accession: B81857  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-289 <PAR>  
 A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CA884858.1; PID:g7380091  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: NMA1630

alignment\_scores:

Quality	Ratio	Length	Gaps
202.00	1.195	288	10
Percent Similarity: 58.681		Percent Identity: 26.389	

alignment\_block:

US-09-303-518D-571 x B81857 ..  
 Align seg 1/1 to: B81857 from: 1 to: 289

67 CTGCTCAATGCTCTCCCTGCTGCTGCTCTCTGTCACACGCTGGG 116

```

7 VALLEUTYValleuGlInPheLeuProHeAlaLeuLeuHsLysLeuAl 23
117 AAACCGGCTCGACATCTGGCGTTTACTTTTAAAGAACAGCCGCC 166
23 aaspleuthrGlyLeuAlaTyrLeuValLysProArgrArgrArgr 40
167 GCATCGTCGCAATATGCGGAGGCGGTTGAACCCCGACAGCAGACG 216
40 leGlyLutLeuLeuAlaLysCysPheProGluTyrPaspGlyLys 56
217 GTCAAAAGCCGTT.....TTTGGGAAACGGCAATGGCGTTT 254
57 ArGlyStHValLeuLysGlnHisPheLysHsMetAlaLysLeuMe 73
255 GGAACCTGGCCCGCGTTTTCAAAACCGAGACATCGAACAATGT 304
73 uGluTyrGlyLeuTyrTrpTyrAlaProAlaGlyArGlyLeuLys 90
305 TCAAAAGCGTACAGCGCTGGGAACAGCTGCAGCGCTTGGCAAGGCG 354
90 alArG...TyrArGAsnLysHsTyrLeuAspAspAlaLeuAlaGly 105
355 GAAGGCGTGTGTTTCATACGCGGACATCGGAGCTTACGATTGGCGG 404
106 GluLysValLleLeuLysGlnHisPheThrAlaPheGluMetAla 122
405 ACAGCTATCATGCGAGCAGCTTCGTTCCATCGTACCGCCATGTACAGC 454
122 ITYAlaLeuAsnGlnAspAlaPro.....LeuLleSerMetTyrSerH 137
455 CGCCGAAATCAAGCATAGCAAAATCATGCGAGCGGAGGCGCGCG 504
137 IsGlnLysAsnLysIleLeuAspLusLleLeuLysGlnArGAsnArG 153
505 GGCAAA.....GGCAAAACCGCGCCCAACCGGCTACAAAG 539
154 TyrHsAsnValPheLeuLleGlyArGThr.....GluG 165
540 GGTCAAAACAAATCATCAAGGCCCTCGCGCGCGGAGCAACATCATC 588
165 yLeuArGAlaLeuValLysGlnPheArGlySerSerAlaProPheLeu 182
589 ..CTGCCGACCGACGCTCCCTTCGCGAGGAAGCGCGC.....GCG 627
182 yLeuProAsp.....GlnAspPheGlyArGAsnAspSer 193
628 GTGTGGCGGATTTTGGCAACCTGCATACACATGACATGGCGGCG 677
194 ValPheValAspPheGlyIleArGThrAlaThrLleThrGlyLeuSe 210
678 AAAATTTGGCACAGTCGCAAGGCGTCAAAACCTTTTTCGTCGAGAC 727
210 rArGlleAlaAlaLeuAlaAsnAlaLysValLleProAlaLleProVal 227
728 GCGTGGCGGAGCAAGGCTTCGTTCACATCCGCCCGCTCAAGGCG 777
227 rGlnLalaAspAsnThr...ValThrLeuHisPheTyrProAlaTyrP 242
778 GAATTCAGCGGCAACAAAGCCGAC...GATGGCGCGCTGTCAACCGCA 824
243 SerPheProSerGlnAspAlaGlnAlaAspAlaGlnArGMetAsnArG 259
825 TACCGAATATGGATAGCGCGCTTTCGCGACGAGTATGTTATGTACA 874
259 eLleGluGlnArGValArGlnGlnHisProGlnLntGlyrPheThrLeu 276
875 ACCGCTATATAACG 888
276 ysaTyrPheLysThr 280
seq_name: p1r2:516888

```

## seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-) - Escherichia coli

N:Alternate names: htrb protein

C:Species: Escherichia coli

C:Date: 13-Jan-1995 #sequence,revision 13-Jan-1995 #text\_change 08-Oct-1999

C:Accession: S16888; A42290; C64848

R:Karow, M.; Georgopoulos, C.

Mol. Microbiol. 5, 2285-2292, 1991

A:Title: Sequencing, mutational analysis, and transcriptional regulation of the Esche

A:Reference number: S16888; MUID:92114808

A:Accession: S16888

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-306 &lt;RAR&gt;

A:Cross-references: EMBL:X61000; NID:948956; PIDN:CA43317.1; PID:948957

A:Experimental source: strain K-12

R:ueguchi, C.; Ito, K.

J. Bacteriol. 174, 1454-1461, 1992

A:Title: Multicopy suppression: an approach to understanding intracellular function in

A:Reference number: A42290; MUID:92165719

A:Accession: A42290

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 200-306 &lt;UEG&gt;

A:Cross-references: EMBL:X59939; NID:942028

A:Experimental source: strain K-12

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Zhao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: C64848

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-306 &lt;BLAT&gt;

A:Cross-references: GB:AE00206; GB:U00096; NID:q1787282; PIDN:AACT4138.1; PID:q17872

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:gene: htrb

A:Function: acylates the intermediate (Kdo)2-1-lipid IVA

A:Description: lipopolysaccharide biosynthesis

A:Note: required for bacterial viability above 33 degrees C in rich media

C:Keywords: acyltransferase; lipopolysaccharide biosynthesis; transmembrane protein

F:21-37/Domain: transmembrane #status predicted &lt;TM&gt;

## alignment\_scores:

Quality:	196.50	Length:	284
Ratio:	1.268	Gaps:	8
Percent similarity:	54.577	Percent identity:	26.056

## alignment\_block:

US-09-303-518d-571 x S16888 ..

Align seg 1/1 to: S16888 from: 1 to: 306

```

103 CTGACAGCTGGGAAACCGGCTCGACATCGCGTTTACTTTAA 152
154 .....GAGACCGCGCGCATCTCGCCATATGCGGAGCG 192
70 etsGrlnglunArGArGlysmetValLysAsnPhelGluSerVal 86
193 GGTTCGACCCCGACACGAGCGGTCAAGCGCTTTTGGGAACGCG 242
87 GlymetGlyLeu..... 90

```







## alignment\_block:

US-09-303-518D-571 x B91036

Align seg 1/1 to: B91036 from: 1 to: 328

```

72 CAATGCTCTCCGCTGCTGCTTTCCTGCTGACACAGCGTGGAAACC 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 GlnLeuProTyrProValLeucysPheLeuGlyThrArgIleGlyAlaMe 69
122 GCGTCGACATCTGCGCTTACCTTTAAAGAACCGCGCGGCATC 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 talArg.....ProPheLeuLysArgArgGlnSerIleA 81
172 GTCGC.....CAATGCGCGAGCGGCGGTTT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 laArgLysAsnLeuGlnLeucysPheProGlnHisSerAla..... 94
198 GAACCCGACACGACGACGACGACGACGCTTTTGGCGAAGCGCAAAAT 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 GluGluArgGlu.LysMetIleAlaGlnAsnPheArgSerLeuGlyMetA 111
248 GCGGTTGGAACCTTCCCCCGCTTTTCAAAAACCGGAACATCGAA 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 laLeuValGluThrGlyMetAlaIleThrPheTrpProAspSerArgValArg 127
298 ACAATGTTCAAGGCGTACACGCGTGGGAACACGTGACAGCGCTTGA 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 LysTrpPheAsp...ValGluGlyLeuAspAsnLeuLysArgAlaGlnMe 143
348 CAAGCGCGAAGCGCTGCTGCTTCATCAGCGCGACATCGACCTACGATT 397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
143 tGlnAsnArgGlyValMetValValGlyValHisPheMetSerLeuGluL 160
398 TGGGCGAGCGCTACATCAGCAGCAGCTTCGCTCCACCGCACCGCATG 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 euGlyGlyArgValMetGlyLeuGlyGlnPro.....MetMetAlaThr 174
448 TACAGCGCGCGAAGATCAAGGATGACAAATATCATGACGCGCGCAG 497
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 TyrArgProHisAsnAsnGlnLeuMetGluTrpValGlnThrArgGlyAl 191
498 GGTGCGCGGCAAGGCAAAACCGCGCCACCGCATACAAAGGCGTCAAC 547
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
191 gMetArg.....SerAsnLysAlaMetIleGlyArgAsnAsnLeuArg 206
548 AAATCATCAAGCGCTGCGCGCGCGGCGGACGACATCATCTGCCGAC 597
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 lYleValGlyAlaLeuLysGlyGlnAlaValTrpPheAlaProAsp 222
598 CACGTCCTCTCCGAGGAAGCGCGCTGCGGCGGATTTTTCGG 647
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 Gln.....AspTyrGlyArgLysGlySerSerPheAlaProPheAla 237
648 C...AAACCTGATACACCATGACACTGCGGCAAAATGGCACACGTCA 694
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
237 aValGlnAsnValAlaThrThrAsnGlyThrValLeuSerArgLeuS 254
695 AAGCGGTAAACCTGTTTTCGTGTCGGAAGCGCGCGCGACGCA 744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
254 ergLysAlaAlaMetLeuThrValThrMetValArgLysAlaAspTyrSer 270
745 GCGTTCGTTGACATCGCGCGCGCTCCAAAGGGAATGGAAGCGC 789
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
271 GlyTyrArgLeuPheIleThrPro.....GlnMetGluGlyTyr 284
790 .....ACAAGCGCAACGATGCGCGCGTTCACCGCAATACCG 829
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 oThAspGlnAsnGlnAla.....AlaAlaTyrMetAsnLysIleIle 299
830 AATTTGGATACGCGTTTTCGACGACGATATCTTTATGTTACACCG 879
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 lLysGluIleMetArgAlaProGluGlnTyrLeuTrpIleHisArg 315

```

```

880 TATTAACG 888
   ||| ||| |||
316 PheLysThr 318

```

seq\_name: p1r2:G65011

seq\_documentation\_block:

```

probable lauroyl acyltransferase b2378 - Escherichia coli
N:Alternate names: probable lipid A biosynthesis protein b2378
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999
C:Accession: G65011
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; M01D:97426617
A:Accession: G65011
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-328 <BIAT>
A:Cross-references: GB:AE000326; GB:U00096; NID:91788718; PIDN:AACT5437.1; PID:917887
A:Experimental source: strain K-12, substrain M61655
C:Keywords: acyltransferase; lipopolysaccharide biosynthesis; transmembrane protein
F:20-40/Domain: Transmembrane #status predicted <TM1>
F:125-145/Domain: transmembrane #status predicted <TM2>

```

## alignment\_scores:

Quality:	196.00	Length:	287
Ratio:	1.195	Gaps:	12
Percent Similarity:	57.143	Percent Identity:	27.526

alignment\_block:

US-09-303-518D-571 x G65011

Align seg 1/1 to: G65011 from: 1 to: 328

```

72 CAATGCTCTCCGCTGCTGCTTTCCTGCTGACACAGCGTGGAAACC 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 GlnLeuProTyrProValLeucysPheLeuGlyThrArgIleGlyAlaMe 69
122 GCGTCGACATCTGCGCTTACCTTTAAAGAACCGCGCGCATC 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 talArg.....ProPheLeuLysArgArgGlnSerIleA 81
172 GTCGC.....CAATGCGCGAGCGGCGGTTT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 laArgLysAsnLeuGlnLeucysPheProGlnHisSerAla..... 94
198 GAACCCGACACGACGACGACGACGCTTTTGGCGAAGCGCAAAAT 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 GluGluArgGlu.LysMetIleAlaGlnAsnPheArgSerLeuGlyMetA 111
248 GCGGTTGGAACCTTCCCCCGCTTTTCAAAAACCGGAACATCGAA 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 laLeuValGluThrGlyMetAlaIleThrPheTrpProAspSerArgValArg 127
298 ACAATGTTCAAGGCGTACACGCGTGGGAACACGTGACAGCGCTTGA 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 LysTrpPheAsp...ValGluGlyLeuAspAsnLeuLysArgAlaGlnMe 143
348 CAAGCGCGAAGCGCTGCTGCTTCATCAGCGCGACATCGGACGCTACGATT 397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
143 tGlnAsnArgGlyValMetValValGlyValHisPheMetSerLeuGluL 160
398 TGGGCGAGCGCTACATCAGCAGCAGCTTCGCTCCACCGCACCGCATG 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 euGlyGlyArgValMetGlyLeuGlyGlnPro.....MetMetAlaThr 174
448 TACAGCGCGCGAAGATCAAGGATGACAAATATCATGACGCGCGCAG 497
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

175 TYRARGPROHISASNASNGLNLEUKETCUTRPVALGINTHRARGLYAR 191
498 GGTGGCGCGCAAGCAAAACCGGCCACCGCATACAGGGGTCAAC 547
191 gmetArg.....SerAsnLysAlaMetIleGlyArgAsnAsnLeuArg 206
548 AATATCATAGGCCCTCGCGCGCGGAGGACACCATCATCTCCGCGAC 597
206 LYLLEVALGLALALEUYSGLYGLIALVALTRPHEALPROASP 222
598 CACGTCCCTTCCTCCGAGAGGCGCGCGGTGCGGATTTTCGG 647
223 Gln.....AspTyrGlyArgLysSerSerPheAlaProPheAl 237
648 C...AAACGTGCATACACCATGACATGCGCAAAATTGGACACGTCA 694
237 aValGluAsnValAlaThrThrAsnGlyThrTyrValLeuSerArgLeu 254
695 AAGCGGTGAAGAACCTGTTTCTGCTGCGAAGCGCTGCCGACGACAA 744
254 erGlyAlaAlaMetLeuThrValThrMetValArgLysAlaAspTyrSer 270
745 GCCTTCGTGTTCACATCCGCCCGCCACAGGGGAATTGACGGC.... 789
271 GLYTyrArgLeuPheIleThrPro.....GluMetGluGlyTyrPr 284
790 .....AACAAAGCCACGATGCCCGCGGTGTCAACCGCAATPACG 829
284 oThrAspGluAsnGlnAla.....AlaAlaTyrMetAsnLysIleIleG 299
830 AATATGATGATACGCCGTTTCCGACGACGATATCTGTTATGTACAACGC 879
299 LulysGluIleMetArgAlaProGluGlnTyrLeuTrpIleHisArgArg 315
880 TATAAAGC 888
316 PheLysThr 318

seq_name: plr2:D85880

seq_documentation_block:
probable heat shock protein dgg [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85880
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85880
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-328 <STO>
A:Cross-references: GB:AE005174; NID:q12516752; PIDD:AA657504.1; GSPDB:GN00145; UMGF:236
A:Experimental source: strain O157:H7, substrain EDJ933
C:Genetics:
A:Gene: dgg

alignment_scores:
Quality: 196.00 Length: 287
Ratio: 1.195 Gaps: 12
Percent Similarity: 57.143 Percent Identity: 27.526

alignment_block:
US-09-303-518D-571 x D85880 ..

Align seg 1/1 to: D85880 from: 1 to: 328

72 CAATATGATGATACGCCGTTTCCGACGACGATATCTGTTATGTACAACGC 121
|||.....|||.....|
53 GlnLeuProTyrProValLeuLysPheLeuGlyThrArgIleGlyAlaMet 69

```

```

122 GCGTCGACATCTGCGCTTTACCTTTTAAGGAAGACCGCGGCATC 171
69 TalaArg.....ProPheLeuLysAlaArgGluSerIleAla 81
172 GTCGC.....CAATATGCGGACGCGGGCTT 197
81 laArgLysAsnLeuGluLeuLysPheProGlnHisSerAla..... 94
198 GAACCCGACACGACGACGCTGCAAAACCGTTTGGCAAGCGCAAAAT 247
95 GluGluArgGluLysMetIleAlaGluAsnPheArgSerLeuGlyMet 111
248 GCGGTTTGGAACTGCCCGCGGTTTCAAAAAACGGAAGACATGAA 297
111 LeuValGluThrGlyMetAlaTrpThrProAspSerArgValArg 127
298 ACAATGTTCAAAAGCGGTACACGCTGGACACGTCGACGCGCTTGA 347
128 LysTrpPheAsp...ValGluGlyLeuAspAsnLeuLysAlaGluMet 143
348 CAAGGCGGAAGGCGTGTGTTATCATCAGCCGACATCGGACGTACGAT 397
143 tGlnAsnArgGlyValMetValValGlyValHisPheMetSerLeuGlu 160
398 TGCGCGACGCTACATCAGCAGCAGCTTCGCTTCACCTGACCGCATG 447
160 euGlyLysArgValMetGlyLeuGlyCysGlnPro.....MetMetAlaThr 174
448 TACAAAGCCCGCAAAATCAAGCGATAGACAAATCATGACGAGCGGAC 497
175 TyrArgProHisAsnAsnGlnLeuMetGluTyrPvalGlnThrArgGly 191
498 GGTGCGCGCAAGGCAAAACCGCGCCACCGCATACAGGGGTCAAC 547
191 gmetArg.....SerAsnLysAlaMetIleGlyArgAsnAsnLeuArg 206
548 AATATCATGACGCGCGCGCGCGGAGGACCAACATCATCTGCGCGAC 597
206 LylleValGlyAlaLeuLysLysGlyGlnAlaValTrpPheAlaProAsp 222
598 CACGTCCCTTCCTCCGAGAGGCGCGCGGTGCGGCGGATTTTCGG 647
223 Gln.....AspTyrGlyArgLysGlySerSerPheAlaProPheAl 237
648 C...AAACGTGCATACACCATGACATGCGCGGCAAAATTGGACACGTCA 694
237 aValGluAsnValAlaThrThrAsnGlyThrTyrValLeuSerArgLeu 254
695 AAGCGGTGAAGAACCTGTTTCTGCTGCGAAGCGCTGCCGACGACAA 744
254 erGlyAlaAlaMetLeuThrValThrMetValArgLysAlaAspTyrSer 270
745 GCCTTCGTGTTCACATCCGCCCGCCACAGGGGAATTGACGGC.... 789
271 GLYTyrArgLeuPheIleThrPro.....GluMetGluGlyTyrPr 284
790 .....AACAAAGCCACGATGCCCGCGGTGTTCACCGCAATPACG 829
284 oThrAspGluAsnGlnAla.....AlaAlaTyrMetAsnLysIleIleG 299
830 AATATGATGATACGCCGTTTCCGACGACGATATCTGTTATGTACAACGC 879
299 LulysGluIleMetArgAlaProGluGlnTyrLeuTrpIleHisArgArg 315
880 TATAAAGC 888
316 PheLysThr 318

seq_name: plr2:E82847

seq_documentation_block:
lipid A biosynthesis lauroyl acyltransferase XF0104 [imported] - Xylella fastidiosa (

```

C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_rev1sion 20-Aug-2000 #ext\_change 20-Aug-2000  
C:Accession: E82847  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82815; MIMD:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: E82847  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-279 <SIM>  
A:Cross-references: GB:AEO03864; GB:AEO03849; NID:g9104879; PIDN:AAE82917.1; GSPDB:GN001  
A:Experimental source: strain 945C  
R:Simpson, A.J.G.; Reilach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Brito  
Neto, E.; Docena, C.; El-Deiry, H.; Facinanci, A.P.; Ferreira, A.J.S.  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; From  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A.  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira  
M.; Tshako, M.H.; Vailada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0104

alignment\_scores:  
Quality: 193.00 Length: 270  
Ratio: 1.235 Gaps: 10  
Percent Similarity: 56.667 Percent Identity: 27.407

alignment\_block:  
US-09-303-518D-571 x E82847

Align seg 1/1 to: E82847 from: 1 to: 279

```
112 CTGGGAACCGGCTGGACATCTGGCTTTACCTTTAAAGAACCG 161
|||||: : : : : : : : : : : : : : : : : : : :
15 LGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 31
|: : : : : : : : : : : : : : : : : : : :
162 CGGGGCGATCGTGGCAATATCGGCGAGCGGTTGAACCCGACAC 211
|: : : : : : : : : : : : : : : : : : : :
31 GCGSALATAGTATGATGATGATGATGATGATGATGATGATGAT 48
||| : : : : : : : : : : : : : : : : : : : :
212 AGACGCGTAAAGCCGTTTTCGCGAAGCGAATATGCGT 252
||| : : : : : : : : : : : : : : : : : : : :
48 TATGATGATGATGATGATGATGATGATGATGATGATGATGAT 64
||| : : : : : : : : : : : : : : : : : : : :
253 TTG...GAACCTGGCGCGGCTTTTCAAAAACCGGAGACATGCAAC 299
||| : : : : : : : : : : : : : : : : : : : :
65 LGGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 78
||| : : : : : : : : : : : : : : : : : : : :
300 AATGTTTAAAGCGGTACAC...GGCTGGGAGACATGCAACGAGCTT 343
||| : : : : : : : : : : : : : : : : : : : :
78 TTTGGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 95
||| : : : : : : : : : : : : : : : : : : : :
344 TGGACAGGCGGAGGCGGTGCTTATCAAGCGGCGGAGACATGCAAC 393
||| : : : : : : : : : : : : : : : : : : : :
95 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 111
||| : : : : : : : : : : : : : : : : : : : :
394 GATTTGGCGGCGGCTATACGACGACGACGCTTCCGCTGACGACGCG 443
||| : : : : : : : : : : : : : : : : : : : :
112 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 126
||| : : : : : : : : : : : : : : : : : : : :
444 CATTTACAGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 493
||| : : : : : : : : : : : : : : : : : : : :
126 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 143
||| : : : : : : : : : : : : : : : : : : : :
```

seq\_documentation\_block:  
lipid A biosynthesis [auroy] acyltransferase (EC 2.3.1.-) (heat shock protein) [import  
C:Species: Salmonella enterica subsp. enterica serovar Typh  
A:Note: This species has also been called Salmonella typh  
C:Date: 09-Nov-2001 #sequence\_rev1sion 09-Nov-2001 #ext\_change 27-Nov-2001  
C:Accession: AD0637  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AD0637  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-306 <PAR>  
A:Cross-references: GB:AL51382; PIDN:CAD08279.1; PID:g16502326; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY1192  
C:Keywords: acyltransferase

alignment\_scores:  
Quality: 181.50 Length: 276  
Ratio: 1.134 Gaps: 12  
Percent Similarity: 57.971 Percent Identity: 26.087

alignment\_block:  
US-09-303-518D-571 x AD0637

Align seg 1/1 to: AD0637 from: 1 to: 306

```
103 CTGGACACGCTGGGAACCGGCTGGACATCTGGCTTTTAAAGAACCG 152
```







A:Residues: 1-453 <TEF>  
 A:Cross-references: GB:AE002295; GB:AE002160; NID:97190314; PIDN:AAF39146.1; PID:97190314  
 A:Experimental source: Strain N199 (MoPn)  
 C:Genetics:  
 A:Gene: TC0278

alignment\_scores:  
 Quality: 162.00 Length: 342  
 Ratio: 0.970 Gaps: 13  
 Percent Similarity: 48.830 Percent Identity: 23.977

alignment\_block:  
 US-09-303-518D-571 x F81720 ..

Align seg 1/1 to: F81720 from: 1 to: 453

```

22 CTGTTTCCCTTGGGACCGCCATGCATCTGTTG..... 60
   |||||  |||||  |||||  |||||  |||||  |||||
2 LeuphelysmetleuargserglyLysValleuValaspHisleuVa 18
61 .....ACGCGCTGCTCAATGCTCTCCCTGCTGCTGCTT 97
   |||||  |||||  |||||  |||||  |||||  |||||
18 lTyrglyleuglyleucllylleuthrilleuargyleuProargS 35
98 CCTGCTGCACACGCTGGGAAACCGCTCGACATCTGGCGTTTACTT 147
   |||||  |||||  |||||  |||||  |||||  |||||
35 erserleuglnleupheglyLysglyleucllythrillepHeTyVal 51
148 TTTAAGGAGACCGCGCGCGCATGCTGCGCAATATGCGG..... 186
   |||||  |||||  |||||  |||||  |||||  |||||
52 lIleerasphearglysthrAlaleuthrasnleuAlaleuAlaphepr 68
187 .....CAGCGCGGTTTGAACCCGACACG 211
   |||||  |||||  |||||  |||||  |||||  |||||
68 ogulysserpherthrighuargTyrglnllealaleuLys...SerValg 84
212 AGACGGTAAACCGCTTTTGGGAAACGCGCAAA..... 246
   |||||  |||||  |||||  |||||  |||||  |||||
84 lnglnVallleIlethrheileglnleuAlathrValaspLysPheAla 100
247 .....TGCGTTTGGAACTTGCCCCC.. 267
   |||||  |||||  |||||  |||||  |||||  |||||
101 LysHisIleasprglnllethrillealaserSerghuaspLapProgl 117
268 .GCGTTTTCAAAAAACCGAAGACATGCAACAATGTTCAAAACGGTAC 316
   |||||  |||||  |||||  |||||  |||||  |||||
117 uglyPhephe.....ProgluglnuValserSerghlnnglnleu... 130
317 AGCGCTGGGACACGTCGACAGGCTTTGACAAAGGCGAAGGCTGCTG 366
   |||||  |||||  |||||  |||||  |||||  |||||
131 .....Asnasphepserargyleuaspnglnnglnlyalale 144
367 TTCATCACCCGACATCGGACGTAAGATTGGCGGACGCTACATCAG 416
   |||||  |||||  |||||  |||||  |||||  |||||
145 leuphecyglyHisglnAlasntprglnleuProphelutyrIleth 161
417 CCAAGCAGCTTCGCTTCACCTGACCGGCATGTACAAGCGCGCAAAATCA 466
   |||||  |||||  |||||  |||||  |||||  |||||
161 rlysaTygTyPProgly.....leuAlapheAlaspProvalLysasp 176
467 AAGCGATAGCAAAATCATGACGCGGCGGACGGTGGCGGCAAGGCAAA 516
   |||||  |||||  |||||  |||||  |||||  |||||
176 roargyleuasnarglylleIleaserleuargLuserPheglnnglyLys 192
517 ACCGCGCCACCGGCAATACAGGGGTCAAAACAATATCATCAAGGCGCTGG 566
   |||||  |||||  |||||  |||||  |||||  |||||
193 lIleValpProProgl..AsnAlatleasnnglnAlaleuargAlaleuH 208
567 CGGCGGAGGACCAACCATCTCTGCGCGACGATGCTCTTCCGACAG 616
   |||||  |||||  |||||  |||||  |||||  |||||
208 slYsglyglnuValalglylleValalaspnglnValleuLeuSerSer 225

```

```

617 AAGCGCGCGCGTGTGGCGGATTTTTCGCAAAACCTGCATACACCATG 666
   |||||  |||||  |||||  |||||  |||||  |||||
225 lnyrSer.....TyrProleupheglyserglnAlapheThr 238
667 ACACTGGCGCAAAATTTGGACATCAAGGCTGCAAAACCTGTTT 716
   |||||  |||||  |||||  |||||  |||||  |||||
239 ThrserProAlaleuAlatYrlyThrargLysProvalIleAlaVa 255
717 CTGCTGGCAACCGCTGCCGACGCAAGGCTGTGTCGATCCGCC 766
   |||||  |||||  |||||  |||||  |||||  |||||
255 lAlaleTyrrarglnProasnnglyasnTyrrleuVal..... 268
767 CCGTCAAGGCGAATGTGACGCGCAACAAGCCACGATGCCCGGTTC 816
   |||||  |||||  |||||  |||||  |||||  |||||
269 .....ProserLysAlaPhe 273
817 AACCGCATACCGAATATGTGATACG..... 843
   |||||  |||||  |||||  |||||  |||||  |||||
274 TyralaasnThrghleuSerIleargLuserThrghlnleuMetas 290
844 .....CGTTT.....CCGACGCACT 859
290 pLysleuMetargheleuLysglylIlealacysLysProgluInT 307
860 ATCTGTTATGTACACCGCTATATA 885
   |||||  |||||  |||||  |||||  |||||  |||||
307 rpleuThrpleuHisLysargTyrrLys 315
seq_name: p1R:E71631

```

seq\_documentation\_block:  
 lipid A biosynthesis lauroyl acyltransferase (htrb) RP718 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C:Date: 21-Nov-1998 #sequence,revision 21-Nov-1998 #text,change 03-Nov-2000  
 R:Accession: E71631  
 R:Andersson, S.G.E.; Zomrodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark  
 Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A71650; MUID:99039499  
 A:Accession: E71631  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-290 <AND>  
 A:Cross-references: GB:AJ235273; GB:AJ235269; NID:93861237; PIDN:CAA15149.1; PID:e134  
 A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: htrb; RP718

alignment\_scores:  
 Quality: 159.00 Length: 205  
 Ratio: 1.214 Gaps: 9  
 Percent Similarity: 63.902 Percent Identity: 21.951

alignment\_block:  
 US-09-303-518D-571 x E71631 ..

Align seg 1/1 to: E71631 from: 1 to: 290

```

289 GACATCGAACAATGTTCAAGCGGTACACGCTGGGACACGTCGACGA 338
   |||||  |||||  |||||  |||||  |||||  |||||
95 gluleuGluserargyleglnlleIle...glyLysgluasnIleuLys 110
339 GCGTTTGACAAAGGCGAAGGCTGTTCATCACCACCGCACATCGCA 388
   |||||  |||||  |||||  |||||  |||||  |||||
110 sleuGlusp.....GlnProphelLeuPheSerGlyHisPheAla 125
389 GCTACGATTGCGGCGAGCATACACGACGAGCTTCGTTCCACCTG 438
   |||||  |||||  |||||  |||||  |||||  |||||
125 snTpraspIleSerleuLysValleuLysLysTyrrTyPPro...LysVal 140
439 ACCGCGATGTACAAGCCCGCAAAATCAAGCGATGTACAATAATCATGA 488
   |||||  |||||  |||||  |||||  |||||  |||||

```

```

                                Quality: 156.50      Length: 273
                                Ratio: 1.057        Gaps: 10
                                Percent Similarity: 54.212      Percent Identity: 23.810

                                alignment_block:
                                S-09-303-518D-571 x T42129 ..

                                sign seq 1/1 to: T42129 from: 1 to: 343

121 GGGCGGCGACATCTGGCGCTTTTACCTTTTAAAGAGACGCGCGCGCAT 170
    :|||:||||| |||:||||| ||| |||:||||| ||| |||
80 LysLeuGlyMetLeuValaGlyArgLeuGlyLysSerLeuArgGlnArgAl 96
171 CGTGGCGCAATATGGCGG.....CAGCGGGCTTTGAACCCCGACA 208
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
96 aLeuIleasnLeuSerLeuGlySerPheArgIleuTygSerAspLysGlyLys 113
209 CGCAGACGGTCAAAACCGCTTTTGGCGAAAGCGCAAAATCGGCTTTGGA 258
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:|||||
113 LysnIleValAspAlaMetPheAlaThrLysSerMetAlaValaValLeu 129
259 CTTGGCCCGCGCGCTTTTCAAAAACCGGAAAGCATGCAACATGTTTCA 308
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
130 MetAlaGlyLeuValaLeuSerGlyProArgLysLysSerHisArgIleAl 146
309 AGCGGTACACGCGCTGGGAACACGTGCAGCAGCGCTTTGGACAAGGCGAAG 358
    :|||:||||| ||| |||:||||| ||| |||:||||| ||| |||:|||||
146 g...TTrasnGlyLeuGlnIleValaGlyLysMetAlaGlnAsnGlnL 162
359 GCGCTGCTGTTCATCAGCGCGCAACATCGGCGAGCTACGAGTATTCG 399
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:|||||
162 yValIlePheLeuValaProHisAlaThrGlyValaAspIleProAlaMet 178
400 .....GGCGAGCGTACATCAAGCAGCAGCTTCGTTCCACCT 437
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
179 LeuMetAlaIleSerGlyArg.....LysMe 187
438 GACCGCATATGTAACAAGCGCGCAAAATCAAAAGCGATACACAAATCATGC 487
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:|||||
187 tAlaIleMetPheHisAsnGlnArgAsnProValaValAspTyrValaIle 204
488 AGCGGCGCAGGCTGGCGCGCAAGGCAAAACCGCGCCACCGCATACAA 537
    :|||:||||| ||| |||:||||| ||| |||:||||| ||| |||:|||||
204 snSerValaTygArgThrPheGlyLysLysLeuHisAlaArgAsn...Asp 219
538 GGGGTCAACAAATATCATAGCGCGCTGGCGCGCGCGCGCAACCATCAT 587
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:|||||
220 GlyIleAlaSerPheValaArgSerValaArgGlnIleuTygTrpGlyTyTy 236
588 CCGTGGCC.....GACCAAGTCCCTTCTCGCGCAAGAGCGCGCGCTGT 631
    :|||:||||| |||:||||| |||:||||| ||| |||:||||| ||| |||:|||||
236 TrpGlyTrasnArgLysAsnArgLysArgProGlnPheSerGlu.....P 249
632 GGGCGGATTTTGGGCAACCTGCAATACACCATGACATGACATGCGCGCAAA 681
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:||||| ||| |||:|||||
249 heAlaAspPhePheHisAlaThrTygLysAlaThrLeuProValaIleGlyArg 265
682 TTGGCACAACGTCAAAAGCGGTGA.....ACCTGCTTTTCTGCTGGCA 725
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:||||| ||| |||:|||||
266 LeuSerArgLysSerGlyAlaArgIleIleProLeuPhe..... 278
726 AGCGTCCCGCGCGACAA.....GGCTTCGCTGTTCACATCGCGCGCG 769
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:||||| ||| |||:|||||
279 .ProValaTygAspArgLysLysHisHisLysLeuThrIleHisValaSerPro 295
770 TCCAAGGGCAATGAACGCGCAAAAGCGACAGATGCGCGCGCTTTGCAAC 819
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:||||| ||| |||:|||||
295 roLeuAlaIleArgGlnLysSerAspArgHisLysIleAlaArgGlnIleasn 311
820 CGCAATAGCAATATGAGTACGCGCTTTTCCGACGCAATATCTGTATT 869
    :|||:||||| |||:||||| ||| |||:||||| ||| |||:||||| ||| |||:|||||
112 GluValaValaGlnAsnPheValaArgProHisArgProGlnIleuTygThrTrpI 328

```

870 GTACACCGCTATTAACG 888  
 : : : : :  
 328 eLeuylsLeuLeuylsThr 334

seq\_name: p1r2:T45134

seq\_documentation\_block:  
 hypothetical protein [imported] - Microbacterium ammoniaphilum (fragment)

C:Species: Microbacterium ammoniaphilum

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000

C:Accession: T45134

R:Strleel, H.M.; Seeber, S.; Jarsch, M.; Kessler, C.

Gene 172, 41-46, 1996

A:Title: Cloning and characterization of the Mami restriction-modification system from

A:Reference number: 222923; MUID:96257250

A:Accession: T45134

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <STR>

A:Cross-references: EMBL:X79027; NID:9984667; PIDN:CAA55649.1; PID:g1679831

A:Experimental source: ATCC 15354

alignment\_scores:  
 Quality: 146.00 Length: 320  
 Ratio: 1.090 Gaps: 19  
 Percent Similarity: 41.875 Percent Identity: 27.188

alignment\_block:  
 US-09-303-518D-571 x T45134 ..

Align seg 1/1 to: T45134 from: 1 to: 529

```

30 CCCCTTTCGACACCGCCATGCTGTTGACCGCCCTGCTCAATGCG 79
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
74 ProAlaGlnArgAlaArgGlnValHisProLeuGlnArgProHisGlnLeu 90
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
80 T.....CTCCCTGCTGCGCTTCTGCTGCACACGCTG 114
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
90 OGlyAlaArgGlyGlyLeuPro...ValArgAspArgAlaGlyArg 106
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
115 GAAA.....CCGGCTCGGACATCT..... 134
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
106 LysArgGlyGlnProAlaArgProAlaAlaArgAspAlaArgHisLeu 122
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
135 .....GGCGTTTACCTTTT...AAAGAAACCGCGCGCGATGCTG 175
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
123 ArgLeuGlyAlaHisProAlaArgAlaGlyValLeuArgHisAlaArg 139
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
176 CCAATATGCGGAGGC.....GGTTTGAACCCGCA..... 206
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
139 GHisArgAlaArgGlyGlyArgGlyArgGlyProArgGlnValProA 156
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
207 .....CAGCGACGCGTCAAGCGCTTTTTCGGAACGCGCAAAATGCG 251
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
156 TglnHisProArgGlyArgArgAspArgAlaGly..... 167
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
252 TTTGCAACTGCGCGCGCTTTTCAAAAACGGAAGACATCGAACA 301
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
168 .....ArgProGlyLeuHisArgArg 175
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
302 TGTTCAAAGCGGTACACGGCTGGAACAGCTGACAGCTTTGACAC 351
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
175 GATGATGAlaArgArgGlnGlyArgProGlnValArgHisGlyArg 192
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
352 GCGGAAGGCGCTGTTTCATCACCGCGCACATCGGCGACCTGATTTGG 401
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
192 spGln.....HisArgAlaAspProArgArgProArgProArgAsp 204
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
402 CGACGCTGATGACGACGAGCTCGTTCCACCTGACCGCGCATGTACA 451
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
205 ArgAlaGlyHisHisProLeuArg.....G 213

```

```

452 AGCGCGCAAAATCAAGCGATGACAAATCATGCA...GGCGGCGACG 498
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
213 uGlyGlyGlnGlyGlnGlnValArgProLeuArgAlaArgGlyGlnG 230
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
499 GT..... 500
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
230 LysGlyProGlyGlyProAlaGlyArgHisProAlaLeuGlyAspGly 246
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
501 .....GGCGGCAAGCAAAACCGCGCGCGCGG.....CA 532
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
247 ArgProArgAlaArgGlnGlyAlaArgAlaAlaHisArgGlyAlaLeu 263
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
533 TACAAAGGCGTCAACCAATCATCAAGCGCCCTGCGCGGCGGCGAGCA 582
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
263 sLeuArgLeuGlnArg.....ArgGlyArgAlaAsp 275
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
583 ATCATCTGCGCCGACACCGCTCTTCCGAGAAAGCGCGCGCGTGG 632
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
275 LysArgAlaGlnGlyArg.....AlaArgArgAlaArgArgAla 288
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
633 GCGGATTTTTCGCAACCTGCATACCATGACATCGCGCGCAAAAT 682
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
289 GlyGly.....GlyGly..... 292
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
683 TGGCACACGTCAAAGCGTGAACCCCTGTTTCTGCTCGCAACCTG 732
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
293 .....LeuProA 295
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
733 CCCGACGCAAGCGCTGCTGTCACATCGCGCCGTCGAAGGGA... 779
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
295 rGArgGlnAspArgValArg...AlaHisArgProArgProArgGly 310
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
780 .....ATTGACGGCAACCAAGCCCATGCGCGCGCTGT 814
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
311 GlyGlyAlaAlaArgLeuAspArgAlaGlyGlyValGlyAlaArgPro 327
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
815 TCAACCGCAA 824
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
327 ArgProArg 330
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq\_name: p1r2:J00405

seq\_documentation\_block:  
 hypothetical 119.5K protein (uvra region) - Micrococcus luteus

N:Alternate names: ORF 1 protein

C:Species: Micrococcus luteus, Micrococcus lysodeikticus

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 04-Feb-2000

C:Accession: J00405

R:Shiota, S.; Nakayama, H.

Mol. Gen. Genet. 217, 332-340, 1989

A:Title: Micrococcus luteus homolog of the Escherichia coli uvra gene: identification

A:Reference number: S04781; MUID:89364717

A:Accession: J00405

A:Molecule type: DNA

A:Residues: 1-1106 <SH1>

A:Cross-references: EMBL:X15867

A:Note: this reading frame extends between two stop codons and does not begin with a

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

alignment\_scores:  
 Quality: 145.00 Length: 366  
 Ratio: 1.007 Gaps: 19  
 Percent Similarity: 39.344 Percent Identity: 25.683

alignment\_block:  
 US-09-303-518D-571 x J00405 ..

Align seg 1/1 to: J00405 from: 1 to: 1106

```

18 CAGGCTGTTCCCTTTCGAAACGCG.....CA 46
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

212 GlnProAlaAspProAlaAlaAspArgLysProAlaArgGlyAlaAlaArg 228
47 TGCACATCCT ..... 56
228 gAlaHisProLeuProGlyAlaAlaArgGlyProAlaArgProGlnGlyA 245
57 .. GTTGAACCGCCCTGCTCAAAATGCTCTGCTGCTGCTGCTGCTGCT 104
245 rGValaArgGlyProValaGlnGlyProVala ..... 254
105 GCACACGCTGGGGAACCGCGCTCGACATCTGCGCTTTACCTTTTAAAG 154
225 .. HIsAlaGlyLeuAlaArgArgArgGlyArgGlyAspArgProAlaLeuG 170
155 AAGACCGCGCGG.....CATGCGCGCAATATG 183
270 yProAlaArgAlaGlnGlnAlaGlnGlnAlaHisAlaArgArgArgG 287
184 CGCGACGGGGGTTTGAAACCGGACACGACGACGTCMAAGCCGTTTTCG 233
287 Ly.....ProAlaArgHisGlnGlnGlyHis..... 295
234 GGAACGCGCAAAATGCGTTTGAATCTGCCCCGCTTTTCAAAAC 283
296 .....ProAlaArgHisArgLe 302
284 CGGAACATCGAACAATCTTCAAGCGGTACGCGTGGGAACACGTG 333
302 uGlyGlyAspArgProGlnGlnGlyArgArgArgGly...ArgV 318
334 CAGCAGCCTTTGACAAGCGCGAAGGCGTGT.....GTATATCAGCG 377
318 aArgGlyArgGlyAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 334
378 GCACATCGGACGCTACGATTTGGCGG..... 404
335 ValaArgArgAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 351
404 ..... 404
351 gGlnAlaLeuLeuProGlnAlaArgAlaAspArgGlyArgAspArg 368
405 .....ACGCTACATCAGCCAGCAGCTTCGTTCCA...CCTGAC 440
368 lAlaLeuValLeuLeuGlnGlnProValaArgArgValProArgValHis 384
441 CGCGATGTACAAAGCCGCGAATAATCAAGCGATGACAAATTCATGCAAG 490
385 ArgHis.....ArgLeuProAlaGln 392
491 CGGCGACGGGTGCGCGCAAGCAAAACCGCGCCGCGCATGCAAGGG 540
392 yGlyProGlyProAlaArgArgArgGlnArgArgAla...ValProAlaArgGly 408
541 GTCAACAATCATC..... 555
409 ArgArgArgAlaValAlaAlaArgGlnValHisLeuGlyLeuAlaAl 425
556 ..AAGCCCTTGGCGCGCGGCGAGCAACCATCATCTGCGCGACAGCTCC 604
425 aArgAlaArgArgAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 442
605 CTTCCTCCGAGAAAGCGCGCGCGCTG..... 632
442 aGlnGlyProAlaAspGlyGlyAlaArgArgArgProAlaArgGlnGly 458
633 .....GGCGATTTTTCGCAAACTGC.....AT 658
459 LeuGlnGlyGlyAlaAspValProGlnProValaArgProAlaArg 475
659 ACACCATGACACTGGCGCAAAATTTGACACAGT.....CAAGGCGG 702
475 uHisHisGlyLeuArgGlyArgHisProLeuArgAspAlaGlnAlaArg 492

```

```

703 AAACCCCTTTTCTGCTGGAGACGCTTGGCGAGCAGG...CTT 749
492 LyAspArgVal.....GlyArgArgProArgAlaLeu 502
750 CGTGTTCACAT.....CCGCCGCTCCAAAGGGAATGGA 784
503 ArgValaHisAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 519
785 ACGGCAACAAAGCCGCGATGCGCGCTGTTTCAACCGCAATACCGA 830
519 oGlnProHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 534

seq_name: p1r2:C97836
seq_documentation_block:
  hypothetical protein htrb [imported] - Rickettsia conorii (strain Malish 7)
  C:Species: Rickettsia conorii
  C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
  C/Accession: C97836
  R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
  Science 293, 2093-2098, 2001
  A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
  A:Reference number: A97700; MUID:21442074; PMID:11557893
  A:Accession: C97836
  A>Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-290 <KUR>
  A:Cross-references: GB:AE006914; PIDN:AL03629.1; PID:g15620214; GSPDB:GN00173
  A:Gene: htrb

alignment_scores:
  Quality: 144.50      Length: 198
  Ratio: 1.194
  Percent Similarity: 61.111      Percent Identity: 21.717

alignment_block:
  US-09-303-518d-571 x C97836 ..

Align seg 1/1 to: C97836 from: 1 to: 290

310 CGGATACAGCGCTGGGAACAGCTGACGAGCTTTGCAAGGCGGAGG 359
||||| ||| |||:|||||:|||||: |||:|||||:
101 AlallelleGlylleGlnHisLeuHisLeu.....GlnGlnGlnPr 115
360 GCTGCTTTTCATACGCGCGCATGCGCAGCTACGATTTGGCGGAGGCT 409
:|||||:|||||: |||:|||||:|||||: |||:|||||:
115 orHeuLeuLeuPheGlnGlnHisPheAlaAsnTrpAspIleSerlleYST 132
410 ACATGAGCGACAGCTTCCGCTTCCAGCTGACGCGCATGTCAAAGCCCG 459
:|||||: ||| |||:|||||:|||||: |||:|||||:
132 LeLeuHisLysPheTrpPro...LysValAlaValIleTyArgGlySAla 147
460 AAATGCAAGCGATAGACAATAATCATGACGCGCGGCGGCTGCGCGCA 509
:|||||: ||| |||:|||||:|||||: |||:|||||:
148 AsnAsnProTyValaHisLysLeuValaHisGlnSerTrpAlaGlyAsp 164
510 AGGCAAAACCGCGCGCACCGCATACAAAGGCTCAACAATCATGCAAG 559
:|||||: ||| |||:|||||: |||:|||||: |||:|||||:
164 sLeuAlaGlnLeuLeuProLysGlyProGlnGlnSerArgAlaLeuValaArg 181
560 CCTGCGCGCGGCGGAGCAACCATCATCTGCGCGACAGCTGCTTCT 609
|||||: ||| |||:|||||: |||:|||||: |||:|||||:
181 lallelYsGlnSerGlnSerlleValaMetLeuValaSpGln..... 194
610 CGCGAGGAAGCGCGCGCTGTGCGGATTTTTCGCAAACTGACATA 659
:|||||: ||| |||:|||||: |||:|||||: |||:|||||:
195 ..LysMetAsnAspGlylleGlnValaProPheLeuGlnHisProAla 210
660 CACCATGACACTGCGCGCAAAATTTGACACAGTCAAGGCGTGAAGCC 709
||| :|||:|||||:|||||: |||

```

A; Scale: Coordination;  
C; Function;





[illegible]

A:Residues: 1-307 <KUR>  
 A:Cross-references: GB:AE008917; PIDN:AAL52296.1; PID:g17983087; GSPDB:GM00190  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BME1115  
 A:Map position: 1  
 C:Keywords: acyltransferase

alignment\_scores:  
 Quality: 137.00 Length: 334  
 Ratio: 0.825 Gaps: 14  
 Percent Similarity: 49.701 Percent Identity: 22.754

alignment\_block:  
 US-09-303-518D-571 x AE3391 ..

Align seg 1/1 to: AE3391 from: 1 to: 307

```

1 ATGTTGCTTACAAATTCAGGCTGTTCCCTTGGCA..... 39
1 MetPheLysLeuLysLeuLysLeuPheArgTrpSerArgLysLeuYsgL 17
40 ..... ACCGCCATGCACATCTCTGTGACGCCGC 67
17 nPheAsnTYRTrpLeuTrpAlaGlnAlaValPheValLeuLeuGlyPheL 34
68 TGCCTCAATGCTCTCCCTGCTGCTGCTTCCGTGTCACACGCTGGCA 117
34 euArgLeuPheProAlaLysAlaAlaIleSerPheSerAlaArgValAla 50
118 AACCGGCTCGACATCTGGCGTTTACCTTTTAAAGAACGCGCGCG 167
51 ArgLeuIleGlyProLeuThr.....ProArgHisLys 61
168 CATGCTGGCC...AATATGGCGGAGCGGCTTGAACCCGACACAGCAGA 214
61 SValAlaThrAspAsnLeuArgLysAla.....TYRProGluLysSerG 76
215 CGGTCAAGCGCTT..... 228
76 LuAlaGluIleValLysIleAlaArgGluMetTrpAspSerMetAlaArg 92
229 ...TTTCCGGAACGCGCAAAATCGGCTTGGAACTTGCCTCCGCGTTT 275
93 LeuPheAlaGluTYR.....IlePheLeuAspAlaValPheAs 105
276 CAAAAACCGGAAGACATCGAACAATGTTCAACGCGTACAGCGCTGGG 325
105 pPheAspProTYRAlaValLysProGlyLeuValGluValGluGlyLeuP 122
326 AACAGCTGACGAGCTTTTGACAGAGCGGCGTGTGCTCATCAGC 375
122 rollePheGluGlyLeuArgAspGluLysLysProHisIlePhePheThr 138
376 CCGCAGATCGGAGCTACGATTTGGCGGCGCTACATCAGCCAGCAGCT 425
139 AlaHisThrGlyAsnPheGluLeu.....Le 147
426 TCCG.....TTCCACCTGACCGCGCATGTACAGC 454
147 uprolleCysAlaAlaThrPheGlyLeuAsnValThrAlaLeuPheArgP 164
455 CGCCGAAATCAAGCGATA...GACAAATATGACGCGGCGGAGGCTG 501
164 roProAsnAsnProTYRlleAlaAsnLysValLeuLysAlaArgAlaGly 180
502 CGGCGAAAGCAAAACCGCGCGCGGATCAAGAGGCTCAACAAT 551
181 ...AsnMetCylHisLeuValProSerLysAlaGlyAlaAlaTrpAla 196
552 CATCAAGCGCTGCGCGGCGGAGGACCATCATCTGCCGAGCAGCAG 601

```

```

196 uAlaGlyIleLeuGlyAspGlyLysAsnValGlyMetLeuValAspGln 212
602 TCCTCTTCCTCCGACGAGGCGCGGCTGTGGCGATTCTTTCGGCAAA 651
213 .....LysPheSerArgGlyValProSerThrPhePheAsnArg 225
652 CCGCATACACCATGACACTGGCGGCAAAATGGCAGCAGCTCAAGGCT 701
226 ProValLysThrAsnProLeuLeuAlaLysLeuAlaArgGlnTYRAspCY 242
702 GAAACCTGTTTCTGCTGCGAACGCTGCCGCGACAGGAGCTTCG 751
242 sasPValTYRProAlaArgCysIleArgLeuProGlyGlyArg...TYR 258
752 TGTTCGACATC.....CGCCCGCTCCAGGGGAA 780
258 rglLeuLysLeuTYRgluArgMetGluLeuProArgAspLysSerGly 274
781 TTGAACGGCAACAAAGCCACGATGCCGCGCTTCAACGCAATACGA 830
275 IleAspIleAsnAlaThr.....ThrGluLeuLeuAsnAspThrValG 289
831 ATATTTGATACGCGCTTTTCGACGCGAGTATCTGTTATGTACACCGCT 880
289 uGlnTrpValArgGluTYRProGlyGlnTrpMetTrpPheHisLysArgT 306
881 AT 882
306 rp 306
seq_name: p1r2:E81551

```

seq\_documentation\_block:  
 lipid A biosynthesis lauroyl acyltransferase, probable CP0676 [imported] - Chlamydia  
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: E81551  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
 A:Reference number: AB1500; MUID:20150255  
 A:Accession: E81551  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-462 <REA>  
 A:Cross-references: GB:AE002225; GB:AE002161; MID:g7189583; PIDN:AAF38487.1; PID:g718  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: CP0676

alignment\_scores:  
 Quality: 137.00 Length: 338  
 Ratio: 0.867 Gaps: 15  
 Percent Similarity: 46.746 Percent Identity: 23.964

alignment\_block:  
 US-09-303-518D-571 x E81551 ..

Align seg 1/1 to: E81551 from: 1 to: 462

```

52 ATCTGTTGACCGCCTGCTCAAAATGCTTCCCTGCTGCTGCTTCTGCTG 101
12 lleLeuGluAlaProLeuTYRtyrLeuValSerGlyIlelleAlaLeuCY 28
102 TCTGCACACG.....CTGGGAAACGCGCTCGAC 130
28 sargHisThrProArgSerPheLeuThrGlyLeuGlyLysGlyPheGlyP 45
131 ATCTGGCGTTTACCTTTTAAAGAACGCGCGCGGATGTGCGCAAT 180
45 heLeuAlaPheTYRlleleSerAspTYRArgLysTYRAlaLeuThrAsn 61

```

```

181 ATGGCGGAGCGGGTTTGACCCGACGACGACGCTCAAGCCGTTT 230
    |||
62 LeuAlaLeuAla.....PheProGluLysThrPh 71
231 TGGGAAAGCGCAAAATGCGT..... 252
    |||
71 eAspGluArgLysIleAlaArgLysLeuGlnHisLeuIleIer 88
253 .....TTGGAAGTTGCC..... 264
    |||
88 hLeuLeuGlnLeuLeuAlaIleGluGlnLeuValGlyAsnIleAspLys 104
265 .....CCGCGGTTTTTCAAAAAACC 284
    |||
105 LeuIleThrIleValThrSerSerArgAsnProLysGlyPheSerSerI 121
285 G.....GAAGACATGCAAAACATGTTCAAAAGCGGTACACG 319
    |||
121 uGluValIleSerAsnGluAspLeuGluGlnThrPheLys..... 134
320 GCTGGGAAACGCTGCAGCAGCTTTGACAAAGGCGAAGGCGTCTGTTTC 369
    |||
135 .....AsnLeuGlnGluLysGlnGlyLeuLeu 144
370 ATCAGCGCGACATGCGACGCTAGCATTTGGCGGACGCTACATGACCA 419
    |||
145 PheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrlleThy 161
420 GAGGCTTCGCTCCACCTGACCGCCATGTACAAAGCCCGCAAAATCAAG 469
    |||
161 sAsnTrpProGly.....IleAlaPheAlaLysAlaIleLysAsnGln 176
470 CGATGACAAATCATGACGCGGCGGCGGCGGCAAAAGCAAAAC 519
    |||
176 rGluSerLysLysIlePheAlaLeuArgLysAlaPheLysGlyLysIle 192
520 GCGGCCACCGGATACAAAGGGGTCAAAACATATCAAGGCGCTGCGCGC 569
    |||
193 ValProProLys...AsnGlyIleGlnGlnGlyIleGluAlaLeuAsnG 208
570 GGGCGGAGCAACCATCATCTGCGCGACACGTC..... 603
    |||
208 nGlyLysLeuValGlyIleValGlyAspGlnAlaLeuLeuMetSerSer 225
604 .....CCTTCCGACGAAAGCGGCGGCTGTGGCGGATTTTTC 645
    |||
225 yThrThrPro.....LeuPhe 230
646 GGCACACCTGATACCATGACACTGCGGCAAAATGCGACACGTCGA 695
    |||
231 GlySerProAlaPheThrThrThrSerProAlaLeuAlaTyrlLysTh 247
696 AGCGCGGAAACCTGTTTCTGCTGCGCAAGCCCTGCCGACGACGCAAG 745
    |||
247 rLysPheProValIleAlaValAsnValSerArg.....GlnAlaLysG 262
746 GCTTGTGTTGCATCCGCCGTCACAGGGGAATGGAACGCAACAA 795
    |||
262 LysPhe.....GluValIlePro...SerAlaLysLeuTyrlAsnLys 275
796 GCC.....CAGCATGCGCGCGTTCACACCGCAAT..... 825
    |||
276 SerLeuProMetLysGlnSerValAlaIleLeuMetAspGlnMetMetG 292
826 ....ACCGAATATGGAATACCGCGTTTCCGACGCGATATCTGTTATGT 871
    |||
292 yPheLeuGlnLysGlyIleAlaSerGlnProGluGlnTrpMetTrpIleH 309
872 ACAACCGCTATATA 885
309 IsLysArgTrpLys 313

```

```

seq_name: p12:B72119
seq_documentation_block:
acyltansiferase - Chlamydia pneumoniae (strain CML029)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: B72119
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: B72119
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <ARN>
A:Cross-references: GB:AE001596; GB:AE001363; NID:94376357; PIDN:AND18251.1; PID:9437
A:Experimental source: strain CML029
A:Genetics:
A:Gene: htrB

alignment_scores:
    quality: 137.00    length: 338
    ratio: 0.867      gaps: 15
    percent similarity: 46.746    percent identity: 23.964

alignment_block:
us-09-303-518d-571 x B72119 ..

Align seg 1/1 to: B72119 from: 1 to: 467

52 ATCTGTTGACCGCCCTGCTCAAAATCCCTCCGCTGCTGCTTTCG 101
    |||||
12 rLeuGlnAlaProLeuTyrlLeuValSerGlyIleIleAlaLeu 28
102 TCTGACAGC.....CTGGAAACCGGCTCGGAC 130
    |||||
28 sArgThrProArgSerPheLeuThrGlyLysGlyPheGlyP 45
131 ATCTGCGTTTACCTTTAAAGAACGCGCGCGATCTGCCAAT 180
    |||||
45 hLeuAlaPheTyrlIleIleSerAspTyrlArgLysThrAlaLeuThrAsn 61
181 ATGCGGACGCGGTTTGAACCCCGACGACGACGCTCAAGCCGTTT 230
    |||
62 LeuAlaLeuAla.....PheProGluLysThrPh 71
231 TGGGAAAGCGCAAAATGCGT..... 252
    |||
71 eAspGluArgLysIleAlaArgLysLeuGlnHisLeuIleIer 88
253 .....TTGGAAGTTGCC..... 264
    |||
88 hLeuLeuGlnLeuLeuAlaIleGluGlnLeuValGlyAsnIleAspLys 104
265 .....CCGCGGTTTTTCAAAAAACC 284
    |||
105 LeuIleThrIleValThrSerSerArgAsnProLysGlyPheSerSerI 121
285 G.....GAAGACATGCAAAACATGTTCAAAAGCGGTACACG 319
    |||
121 uGluValIleSerAsnGluAspLeuGluGlnThrPheLys..... 134
320 GCTGGGAAACGCTGCAGCAGCTTTGACAAAGGCGAAGGCGTCTGTTTC 369
    |||
135 .....AsnLeuGlnGluLysGlnGlyLeuLeu 144
370 ATCAGCGCGACATGCGACGCTAGCATTTGGCGGACGCTACATGACCA 419
    |||
145 PheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrlleThy 161
420 GAGGCTTCGCTCCACCTGACCGCCATGTACAAAGCCCGCAAAATCAAG 469
    |||

```

161 SASnTyProGly.....IleAlaPheAlaLysAlaIleLysAsnGlnA 176  
470 GSATGACAAATATCATGCGAGCGGAGGCGCGGCAAGGCAAAACC 519  
176 rGLeuSerLysLysIlePheAlaLeuArgLysValPheLysGlyLysIle 192  
520 GCGCCACCGGCAATCAAGGGGTCAAAACAAATCATCAAGGCGCTCCGCG 569  
193 ValProProls...AsnGlyIleGlnGlnGlyIleGlnAlaLeuAsnGln 208  
570 GGGGAGGCAACATCATCTCCCGGACACCTC..... 603  
208 nGlyLysLeuValGlyIleValGlyAspGlnAlaLeuMetSerSer 225  
604 .....CTTCTCCGAGAGGCGGCGCGCTGTGGCGCATTTTTC 645  
225 yThrTyPro.....LeuPhe 230  
646 GGCAAACTGATACACCATGACACTGGCGCAAAATTTGGCACACTCAA 695  
231 GlySerProAlaPheThrThrThrSerProAlaLeuAlaTyLysTh 247  
696 AGGGGTGAACCCCTGTTTCTCTCTCGAAGCGCTCCGCGAGGACAG 745  
247 rGlyPheProAlaIleAlaValAlaValSerArg.....GlnAlaLysG 262  
746 GCTTCGTGTTGACATCCGCGCGCTCCAGAGGGAATTGAACGCAACAA 795  
262 LysPhe.....GluValIlePro...SerAlaLysLeuTyAlaAsnLys 275  
796 GCC.....CACGATGCCCGCGCTTCAACCCCAAT..... 825  
276 SerLeuProMetLysGlnSerValAlaIleLeuMetAspGlnMetGln 292  
826 .....ACCGAATATGATGATACCGCGCTTTCGAGCGCATCTCTGTTATGT 871  
292 yPheLeuGlnLysGlyIleAlaSerGlnProGlnGlnTrpMetTrpIleH 309  
872 ACAACCGCTATATAA 885  
309 lAspArgTrpLys 313  
seq\_name: p1r2:D97552  
seq\_documentation\_block:  
lipid a biosynthesis lauroyl acyltransferase (htrb) rp178 [imported] - Agrobacterium tum  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: D97552  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Durillo, B.; Goldman,  
A.; Liu, F.; Moilam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2333-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; PMID:11743194  
A:Accession: D97552  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-312 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK87373.1; PID:G15156679; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_2930  
A:Map position: circular chromosome  
alignment\_scores:  
Quality: 136.50 Length: 287  
Ratio: 0.982 Gaps: 14  
Percent Similarity: 48.432 Percent Identity: 22.648  
alignment\_block:  
US-09-303-518d-571 x D97552 ..  
Align seg 1/1 to: D97552 from: 1 to: 312

190 GCGGGTTTGACCCCGACACGCGACAGCTCAAGCCGTTTTCGGAAC 239  
41 AlagIlyleArgAlaThrAspArgLeuAlaArgLeuIleGlyProLysTh 57  
240 GGCAAAATGGCGTTTG.....GACTGTGCGCCCGGCTTTTCAAA 280  
57 rGlyAlaGlnLysLeuMetLeuTyTrpAsnLeuAlaArgAlaPheProGln 74  
281 MACCGAAGCATGCAAAACATGTTCAAGCGGTACACGCTG..... 324  
74 yThrGlnGln.....GluArgLeuAlaIleAlaMetAspSerTrpAlaAsn 89  
325 .....GACACGTC..... 333  
90 MetGlyArgLeuAlaIleGlnTyValPheLeuAspArgLeuPheAspH 106  
333 ..... 333  
106 eaSPProGlnLysAsnGlnProGlyArgIleGlnValGlnGlyThrSer 123  
334 .....CAGCAGCTTGGACAAAGGCGAAGGCGCTGCTTTCATCAGCGG 378  
123 hPheLeuGlnLeuArgAspAsnProArgProPheIleValPheThrAla 139  
379 CACATGCGGAGCTACGATTGGGCGGCGGCTACATCAGCCACACTTC 428  
140 HisSerGlyAsnPheGlnLeu.....LeuTr 148  
429 G.....TTCACCTGACCGCATGTACAAGCGC 457  
148 oValAlaGlySerAlaPheGlyLeuAspValThrValLeuPheArgPro 165  
458 CCAAAATCAAGCAATA.....GACAAATATCATCAGCGGCGGCTGCGC 504  
165 roAsnAsnProTyValAlaAspLysValPheAsnPheArgLysGlnArg 181  
505 GGCAAAAGCAAAACCGCGCGCATACCAAGGGGTCAAAACAAATCAT 554  
182 ...MetGlyAsnLeuValProSerHisAlaGlySerPheAlaLeuAl 197  
555 CAAGGCGCTGCGCGCGGCGGAGCAACCATCTCTGCGCGACAGTCC 604  
197 aArgGlnLeuGlnLysGlyGlyValGlyValLeuValAspGln.... 212  
605 CTTCTCCGAGGAGGCGCGCGCTGTGGCGGATTTTTCGGAACCT 654  
213 .....LysPheGlyLysGlyLeuThrThrLysPhePheGlyLeuGln 226  
655 GCATACACCATGACACTGCGGCAAAATTTGCACACGTCMAAGGGGTGA 704  
227 ValArgThrAsnProLeuLeuAlaLysLeuValArg..... 238  
705 AACCTGTTTTCGCG.....TGCAGCGCTGCGCG 736  
239 ...GlnPheAsnCysAspValTyProAlaArgCysIleArgLeuProA 254  
737 ACGGCAAGCGTTGCGTTGCATCCGCGCG.....GTCAAGAG 777  
254 sPAsnArg...TyArgLeuGlnIleGlnProLysValGlnIleProArg 269  
778 GAATTGAAGCGCAACAA.....GCCACGATGCCGCGTTTCAACCG 821  
270 AspGlnLysGlyAsnValAspIleGlnAlaThrAlaGlnLeuLeuAsnAs 286  
822 CAATACCGAATTTGGATACCGCGTTTTCGACGACGATGCTGTTATGT 871  
286 pLysValGlnSerTrpValArgGlnTyProGlyGlnTrpLeuTrpTyGln 303  
872 ACAACCGCTAT 882  
303 lAspArgTrp 306





```

131 ATCT.....GGCTTTTACCTTTTAAAGAAAGACGGCGGCATC 171
    :|||
26 gGlyValHisArgGlyValLeu.....ArgHisG 36
    :|||
172 GTCGGCATATGCGCGAGCGGCTTGAACCCGA..... 206
    :|||
36 LurArgGlnSerAlaGlyGly...GlnHisArgGlyGlnHisValGln 51
    :|||
207 .....CACCA..... 212
    :|||
52 ValGlnArgAspGlnLeuHisAlaValAlaArgSerArgProGlyVal 68
    :|||
213 .....GACGCTCAACCCGCTTTTGGCAAGCGCAAA 246
    :|||
68 aAlaGlyValArgArgHisGlyArgGlnArgCysGlyAlaGlnAla 85
    :|||
247 TGCAGTT..... 254
    :|||
85 rGArgAlaLeuGlnProArgArgGlnSerArgSerGlyCysIleValGly 101
    :|||
254 ..... 254
    :|||
102 GlnGlnGlnHisProGlyValProGlnGlyLeuLeuProValCysLeuAlr 118
    :|||
255 .....GGAAGCTTCCCGCCGCTT 273
    :|||
118 gGlyGlnArgGlyValArgArgArgHisProGlyProCysProAlaGlnHisG 135
    :|||
274 TTTCAAAAACCGGAGA.....CATGCAACAAATTTCAAGCGGT 314
    :|||
135 LurGlnArgAlaGlyArgAlaGlyLeuArgArgAlaGlyProHisGlnIly 151
    :|||
315 ACACGCGCTGGGAACACATGCGACAGCTTGGACAGCGGCGAGGCTGC 364
    :|||
152 AspHisValGly.....ValAlaGlyHisProGlyGlyValGlnProAl 166
    :|||
365 TGTCTATCAGCGCGACATGCGACATGATTTGGCGGACGCTACATC 414
    :|||
166 aValArgArgValLeuHisArgArg.....GlyGlyArgValHisG 180
    :|||
415 ACCGACGACCTTCGCTTCCA...CTGACCGCGCATGTCAAGCGCGGAA 461
    :|||
180 LyrProArgGlyValAlaValProCysProHisArgArgAlaAlaHisGlyGln 196
    :|||
462 AAT.....CAAGCGATGACAAATCTATGCG 487
    :|||
197 HisLeuProValProGlyLeuGlyLeuGlnProGlyHisGlyHisGly 213
    :|||
488 A.....GGCGGCGAGGCTGCGCGCGGAA 510
    :|||
213 uLeuArgAlaLeuHisArgLeuArgHisArgGlyProGlyArgLeuLeuAlr 230
    :|||
511 GGCAAAACCGCGCGCGCGCATGACAGGCGTCAACAAATCATGACAGG 560
    :|||
230 rGValProGlyProValArgHisHisArgGlyArgLeuLeuHis..... 244
    :|||
561 CTTGCGCGCGCGCGGACCAACATCATCTGCGCGACGCTGCTTCTTC 610
    :|||
245 .....GlnHisGlyGlnAlaArg..... 250
    :|||
611 CGCAGAAAGCGCGCGCGCTGTGGCGGATTTTTCGCAACCTGCAATAC 660
    :|||
251 .....ArgLeuGlnArgGlnAlaGlyGlyValAlaArgGlnArgValAlaValHis 266
    :|||
661 ACATGACACTGGCGCGCAAAATGGCAACATGCAAGCGCTGAAACCGCT 710
    :|||
266 eArgArgHisGlyGlyAspProGlyGlnArgGlnAspLeuGln..... 280
    :|||
711 GTTTTCTGCTGCAAGCGCTGCGCGACGACAGGCTTCTGCTTGCACA 760
    :|||
281 .....ProValProHisGlnProArgArgAlaArgHis 291
    :|||
761 TCCGCCCCGCTCA.....AGGGAAATGTAACGGCAAC 792

```

```

seq_name: p1r2:C75580
seq_documentation_block:
adenine deaminase-related protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: C75580
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.;
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: C75580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <NH2>
A:Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDN:AAF12376.1; PID:9646
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0268
A:Map position: 2

alignment_scores:
Quality: 132.00 Length: 351
Ratio: 0.852 Gaps: 21
Percent Similarity: 44.160 Percent Identity: 26.496

alignment_block:
US-09-303-518D-571 x C75580 ..

Align seg 1/1 to: C75580 from: 1 to: 376

18 CAGCGTGTTCCTCCCTTTGCGAACCGGCATGACATCTGTGACCGCC 67
   :|||
23 ArgAlaAlaAspGlyLeuGlyAsnArgCysArgHis.....ArgPr 36
   :|||
68 TGCTCAATAGCTCTCCCTGCTGCTGCTTCTGCTGCAAGCT.... 113
   :|||
36 cAla...LeuArgGlyAspAspSerAlaArgLeuAlaGlyHisAlaHisG 52
   :|||
114 .....GGGAAACCGCGCTGCGACATCTGCGCTTTTACCTTTAA 152
   :|||
52 lStrHisPheGlyAlaAlaGlyArgGlyGly.....Thr 63
   :|||
153 GGAAGACCGCGCGCGCATGCTGCGCAATATGCGGAGCGGCTTGAAC 202
   :|||
64 GLyArgProAlaAlaGlyProAlaGlyGlyGlnSerGlyAlaGlyGlyLur 80
   :|||
203 CGGA.....CACGCAAGCGTCAAGCGCTTTT 231
   :|||
80 cArgArgValAlaHisProAlaLeuArgSerGlyGlyArgAlaAspArg 97
   :|||
232 GCGGAAACCGCAAAATGCGGCTTGAAGCTTTCGCAAAATAA 281
   :|||
97 LagArgAlaGlySerAlaAlaGlyAlaArgTyArgGlyLurLeuGly 113
   :|||
282 ACCGGAAGACATGCAAAAT.....GTTCAAG 310
   :|||
114 ProAlaArgProArgGlyAspArgProArgAlaSerArgArgLeuArgAl 130
   :|||
311 CGGTACACGCTGGGAAACA...CTGCAAGCGCTTGGACAAAGCGGCA 357
   :|||
130 aAlaAlaArgLeuAlaThrLeuArgGlyAlaGlyAsnLeuArgGlyArg 147

```

```

358 GGGGCGCTGTTTCATCCAGCCGACATCGGAG..... 389
|||||.....
147 rglYser.....AlaAlaAlaArgArGArAlaProAlaArg 160
390 .....CTACGATTTCG 400
161 TTrArGArGArProAlaAlaArgLeuGlyAlaArgGlyHisLeuArgSerAla 177
401 GGGGCGGCTA...CATCAGCCGAGCTTCCTTCACCTGACCCGACATG 447
|||||.....
177 ArGArAlaLeuAlaHisAlaProAlaSer...ArgProAlaArgHisA 193
448 TACAAGCCGCGGAAATCAAGGATGACAAATCATGACAGC..... 491
|||||.....
193 rglAcySerGlyAlaGlySer.....ArgArGArGArProAlaGlyArgArG 207
492 .....GGGCGAGG 499
208 ArGArProLeuAlaAlaArgArGAlaValAlaGlyLeuLeuAlaSerAlaGlyAla 224
500 TGGCGCGCAAG.....CAAAAGCGCGCCGACCGGACATACAA 537
|||||.....
224 GArAlaAlaArgArGHisProGlyHisGlnHisSerAlaArgArGAlaSerG 241
538 GGGGTCAACAAATCATCA.....GGCGCTCGCGCGG 572
|||||.....
241 LysIleProAlaArgArGAlaArgArGlyProAlaArgGlyGlySerGly 257
573 .....CGAGGACACATCATCTGCGCGA.....CC 598
|||||.....
258 AlaArgThrAlaArgArGArGHisArgProGlyArgGlyArgArGlyArg 274
599 ACGTCCCTTC.....TCGGCAG 615
|||||.....
274 CAArgAlaValAlaProAlaLeuArgGlySerAspAspArgProAlaSerG 291
616 GAAGCGCGCGCTGTGGCGGATTTTTCGCAACCTGCATACACCAT 665
|||||.....
291 LysArgGlyArgArGArgProGly.....ProGly 300
666 GACACTGGCGGCAAAATTCG.....ACAGCTCAAGCGGTGAAA 706
|||||.....
301 AspGlyGlyAlaLeuAlaGlyLeuHisLeuAlaLeuSerArgHisH 317
707 CCGGTGTTTCTG.....CTGC 723
|||||.....
317 sProGluPheSerArgProGluArgAspSerGlyAlaGlnAlaAspAla 334
724 GAAGCGCTGCCGAGCAAGGCTTCGTTCGACATCCGCCCTGCCA 773
|||||.....
334 LysArgProAlaAlaArgArGAspGlyValAlaAlaLeuAlaGlyProGlyArgVal 350
774 AGG 776
|||
351 Arg 351

```

seq\_name: pIiz:AG0251

seq\_documentation\_block:

lipid A biosynthesis acyltransferase (EC 2.3.1.-) [imported] - Yersinia pestis (strain C

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AG0251

R:Packhill, J.; Wren, B.W.; Thomson, N.R.; Tiltall, R.W.; Holden, M.T.G.; Prentice, M.B.;

deno-Tarrage, A.M.; Chillingworth, T.; Cronin, A.; Davis, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AG0251

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-320 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90875.1; PID:g15980074; GSPDB:GN00175

C:Genetics:

A:Gene: msbA

C:Keywords: acyltransferase

alignment\_scores:

Quality	Ratio	Percent Similarity	Length	Gaps
130.00	0.897	53.506	271	13

Percent Identity: 23.247

alignment\_block:

us-09-303-518d-571 x AG0251 ..

Align seg 1/1 to: AG0251 from: 1 to: 320

```

124 CTGGACATCGGGCTTTTACCTTTAAAGAACCCGCGGACATCGT 173
|||||.....
52 TLeGlyArgLeuAlaGlyLysPheAlaLysSerAlaArgArGAlaAr 68
174 CGCCAAAT.....ATCGCGAGCGGGTTTGAAACCCGACACCG 211
|||||.....
68 gTLeAsnLeuLeuTyrCysMetProGluLeuProGluSerGlyArgGly 85
212 AGACGGTCAAGCGCTTTTGGGAAACGCAAAATGCGTTTGGAAT 261
|||||.....
85 ILeIleAspGlnMetPheAlaThrAlaAlaGlnProLeuMetMet 101
262 GCGCGCGGCTTTTCAAAAACGGAACATCGAAACATGTTCAAAGC 311
|||||.....
102 AlAGlyLeuCysPheArgAspProLysLysValLeuThr.....Ar 115
312 GGTACAGCGCTGGACACGCTGCAGCAGCTTTTGAC.....A 349
|||||.....
115 gValHis...trp...HisGlyGlnGluIleLeuAspGlyLeuGlnGln 130
350 AGGCGAAAGGCTGCTTTTCATCAGCCGACATCGGACGTACGAT 399
|||||.....
130 InGluArgAsnValIleLeuLeuValProHisAlaTrpSerIleAsp 146
400 GCGGAGCGCTACATCAGCCAGCAG.....CTCGCTTCACCTGACCGCAT 446
|||||.....
147 ProAlaMetLeuLeuAlaGlnGlnGlyPro.....ValAlaGlyMe 161
447 GTACAGCGCGCGCAAAATCAAGCGATGACAAATCATGACGCGGCA 496
|||||.....
161 TrPheHisHisGlnArgAsnProLeuValAspTyrLeuTrpAsnSerAla 178
497 GGGTGGCGGCAAAAGCAAAACCGGCCGACCGGATACAGGGTCAA 546
|||||.....
178 rGlyLeuHisPheGlyArgGlyLeuHisAlaArgGlu...SerGlyIleLys 193
547 CAATCATGCAAGCGCTTCGCGGCGGCGGAGGCAACATCATCTGCGCA 596
|||||.....
194 ProPheIleSerSerValArgGlnGlyPheTrpGlyTyrTyrLeuProAs 210
597 CCAAGTCCCTTCCTCGCAGAGACCGCGCGCTGTCGGGGAATTTTCG 646
|||||.....
210 rGluAspTyrGlyProGlnGln.....SerGluPheValAspPhePhe 225
647 GCAAACCTGATACACCATGACACTGGCGCAAAATTCGGACACGTC... 693
|||||.....
225 LArTrTyrLysAlaThrLeuProAlaIleGlyArgLeuMetLysLys 241
694 ...AAAGCGTGAACCCCTGTTT.....TTTCGTCTGCGCAAGC 728
|||||.....
242 ArgAlaAlaIleValPrometPheProValTyrAsnTyrArgGlyHisAr 258
729 CTGCGCGGACGAGCAAGGCTTCGTTCGACATCGCGCGGTCGCAAGG 778
|||||.....
258 gLeuAsp.....IleTyrIleArgProPrometAspA 269
779 AATTG...AACGCAACAAAGCCGACGATGCGCGGCTTCACCGCAAT 825

```

```

269 spleuAlaAspAlaAspAlaTyrIleAlaArgArgMetAlaInclu 285
      ||| : : : : : ||| : : : : :
826 ACCGAATATTGATAGCGGCTTTCCGACGAGATCTGTTATGACAA 875
      : : : : : : : : : : : : : : : :
286 ValGluLeuLeuValLysProAsnProGluInThrTrpIleLeu 302
      : : : : : : : : : : : : : : : :
876 CCGCTATAAACG 888
      : : : : :
302 spleuLeuLysThr 306

seq_name: p1r2:T14180

seq_documentation_block:
exit protein - Mycobacterium smegmatis
C:Species: Mycobacterium smegmatis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: T14180; T14163
R:Zhu, W.M.; Arceneaux, J.E.L.; Beggs, M.L.; Byers, B.R.; Eisenach, K.D.; Lundrigan, M.D
Mol. Microbiol. 29, 629-639, 1998
A:Title: Exochelin genes in Mycobacterium smegmatis: identification of an ABC transporter
A:Reference number: Z17906; MUID:98385832
A:Accession: T14180
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1122 <ZHU>
A:Cross-references: EMBL:AF034152; NID:g3421057; PID:g3421058; PIDN:AAC32046.1
R:Yu, S.; Fiss, E.; Jacobs Jr., W.R.
J. Bacteriol. 180, 4676-4685, 1998
A:Title: Analysis of the exochelin locus in mycobacterium smegmatis: biosynthesis genes
A:Reference number: Z17898; MUID:98389687
A:Accession: T14163
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 534-1122 <YUS>
A:Cross-references: EMBL:AF027770; NID:g3560502; PID:g3560505; PIDN:AAC82548.1
C:Genetics:
A:Gene: exit
C:Superfamily: Mycobacterium tuberculosis probable ABC transporter RV0194; ATP-binding C

alignment_scores:
Quality: 130.00 Length: 337
Ratio: 0.828 Gaps: 18
Percent Similarity: 46.588 Percent Identity: 26.409

alignment_block:
US-09-303-518D-571 x T14180 ..
Align seg 1/1 to: T14180 from: 1 to: 1122

3 GTTTCCTTACAAATTCAGGCTGTTCCCTTCCGACCGCCATGCACA 52
  ||| ||| : : : : : ||| : : : : :
135 ValValPheGlyLeuGlnLeuThrAlaProMetIleAlaArgArgVal 151
      : : : : : : : : : : : : : : : :
53 TCCTGTTGACCGCCCTGCAATAGCCTGCTGCTGCTGCTGCTGCT 102
  : : : : : : : : : : : : : : : :
151 u.....AspGlnGlnAlaGlnIleely.....A 159
      : : : : : : : : : : : : : : : :
103 CTGCACACCGCTGGAAACCGGCTCGACATCTGCGTTTACCTTTAA 152
  ||| ||| : : : : : ||| : : : : :
159 rGAlaThrAlaLeuAlaThrAspLeuIleSerGlyGlnArgProLeu 175
      : : : : : : : : : : : : : : : :
153 GGAACAGCCGCGCGCATCGTCCCAATATGCGGAGCGGGTTGAAC 202
  ||| : : : : : : : : : : : : : : : :
176 GlyAsnAlaArgAlaGlyGlnArgGlyAlaLeuProCysArgGluPr 192
      : : : : : : : : : : : : : : : :
203 CCGACAGCAGAC..... 215
  | : : : : |
192 oSerTrgAlaAspArgAspValAlaCysArgAlaHisProGluHisPr 209
      : : : : : : : : : : : : : : : :
216 ..GTCGAAAGCCGTTTTCGGAAACGGCAAAATGCGGTTGAACT... 260
      ||| : : : : : ||| : : : : :

```

```

209 euGIyArg.....CysGlyGlyArg.....GlyThrArg 219
261 .....TGCCCCCGCGCTTT 274
220 ArgAspGlyGlyArgGlyArgGlyLeuLeuArgAspProArg...Ph 235
275 TCATAAAACCGA.....AGAC 291
235 eAspAspGlyGlyProValAspHisArgAspArgProArgProValProH 252
292 ATCGAAACATGTTCAAGCGGTACACG.....CTGGACACGTGCA 335
  ||| : : : : : ||| : : : : :
252 IsArgThrValLeuAlaAlaGlyGlyArgAlaGlnLeuLysSerArgGly 268
336 GCAGCTTGGACAGGCGAAGGCGGTGCTGTCATACCGCGCCGACATCG 385
  ||| ||| : : : : : ||| : : : : :
269 ProGlyValGlyGlnProArgTrgSerArgProGluArgProAspProBr 285
386 GCAGCTACATTT...GGCGAGCGCTACATACCGACGCTCCGTTTC 432
285 oleuProArgLeuArgGlyGlyProProArgArgGlyProProValG 302
433 CACCTGACCGCCATGTACAGCCGCGAA.....ATCGAAGC 470
  ||| : : : : : : : : : : : : : : : :
302 LProGlnArgAsnThrArgGlySerGluLysProValValArgArgArg 318
471 GATAGCAAAATCATGACGCGCGGCGAGGTGCGCAAGCGCAAAACGC 520
  ||| : : : : : : : : : : : : : : : :
319 AlaArgArgIleArgArgGlyThrHisGluArgArgProArgCysArgAr 335
521 CCCCCACCGGCAT.....ACAGGGGTCAACCAATC 552
  : : : : : : : : : : : : : : : :
335 gThrHisArgSerAlaLeuArgIleSerAlaSerArgGlyGlnArgHisG 352
553 ATCGAAGC.....CCTGCGCGGCGGAGGC 578
  ||| : : : : : ||| : : : : :
352 LValGlyArgTrpProAlaProArgProAspProAlaGlu...ArgAla 367
579 AACCATTCCTGCGCCGACACGCTCCTTCCGACGAGCGCGCG 628
  : : : : : ||| : : : : : ||| : : : : :
368 ProArgTrpProAlaArgArgTrpProProGlnLeuValGlnArgAs 384
629 TGTGGCGGATTTTTCGCAACCTGCATACACCATGACGCTGGCGGCA 678
  : : : : : : : : : : : : : : : :
384 PaladLysn.....GlnSerGluCysHisArgAsnArgSerAlaAla 399
679 AAATTGGCACACGTCGTAAGCGGTGTAACCTGTTTCTGCTGCGAACG 728
  : : : : : ||| : : : : : ||| : : : : :
399 rGThrGlyLeuArgArgArgTrpThrThrGly..... 410
729 CCGCGCGAGGACAGAGCTTGTGTCACATCCGCCGCTCCAGGG 778
  ||| : : : : : ||| : : : : : ||| : : : : :
411 ..ValArgArgArgArgArgArgProAla...ProArgTrgSerArgPr 425
779 AATTGAAGCGCAACAAAGCCGATGCGCGCTGTCACCGCAATAC 828
  : : : : : : : : : : : : : : : :
425 ocysArgArg.....GlyAlaArgCysGlnProValArgArgThrPr 440
829 GAATATTGGAT 839
  ||| : : : : : ||| : : : : :
440 roThrLeuAsp 443

seq_name: p1r2:A35938

seq_documentation_block:
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347

```

A:Accession: A35938  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-2248 <GAP>  
 A:Cross-References: GB:J029929  
 C:Genetics:  
 A:Gene: GDB: FLG  
 A:Cross-References: GDB:119912; OMTW:135940  
 A:Map position: 1q21-1q21  
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology  
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat  
 F:246-569/Region: filaggrin repeat  
 F:570-893/Region: filaggrin repeat  
 F:1074-1397/Region: filaggrin repeat  
 F:1573-1896/Region: filaggrin repeat

alignment\_scores:  
 Quality: 127.50 Length: 236  
 Ratio: 1.081 Gaps: 10  
 Percent Similarity: 50.000 Percent Identity: 25.000

alignment\_block:  
 US-09-303-518D-571 x A35938 ..

Align seg 1/1 to: A35938 from: 1 to: 2248

```

198 GAACCCGACACGACGACGTCACCGCTTTTTCGCGAAGCGCAAAAT 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
725 Glnsppatgaspsergluglylserglu.....aspsergluar 738
248 GCGGTTTGACCTGCCCCCGCTTTTCAAAAAACGGAAGACATGAA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
738 ghiserclyseralaserarg.....Asnhsarg 749
298 ACAATGTTCAAGCGGTACACGCTGGGACACGTCGACGCGTTTGA 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
749 lysesrlaglgluglinserserarglserarghlsproglserhls 765
348 CAAGCGGACGAGGCTGCTGTCATCACCGCGCA..... 380
766 Aspgluasparglaglglhlsarglinserserlaspserserglne 782
381 .....CATCGCAGCTA.....CGATTGGCGGACGCTAC 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
782 rcllythrghlsthrglnthrserserarglglalalalaserarg 799
412 ATCAGCCAGCAGCTCCGTCACCTGACCGCATGTACAAGCGCGGAA 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
799 lnglglalalargserarglglasparghlsgllyserglhlsln 815
462 AATCAAGCGATGACAAATCATCAGCG.....GGCAGGCTGC 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
816 lnserlaspserserarghlserserllegllyargglglalalase 832
503 GCGGCAAGGCAAAACCGCGCCACCGCATCAAGGGT..... 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
832 rserlavalargasparglglhlsargllyserargllyserglalals 849
543 .....CAACAATCATCA.....GCGCCT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
849 ertspglnglglhlserserlaspserarglserglinserserlase 865
564 GCGCGGGGCGAGCAACCAT.....CATCTGCCCGGACCGCTGC 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
866 glnglglalaglserhlsnglinserserhlsnglinserserthargl 882
605 CTCTCCGAGGAGGAGGCGGCGGTGGCGGATTTTTCGCAACCT 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
882 yarlgserglngllyargsergllyargsergllyserphelietrglnvals 899
655 GCATACCATGACACTGCGGCAAAATTTGCG.....ACA 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

899 ertlnhlsnglglinserserlaspserlaspsergllyargltharglthrs 915  
 690 CGTCAAGGCGTGAACCCGCTTTTCTGCTGGACGCGCTGCCGACG 739  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 916 gllyargglngllyserhlsnglglinserserlaspserarghls 932  
 740 GACAAGGCTGCTGTT.....GCACATCGCCCGCTC 771  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 932 sserlaspserglnglglinserserlaspserlaspsergllysera 949  
 772 CAAGCGGA 779  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 949 rglargglly 951

seq\_name: p1r2:E90949

seq\_documentation\_block:  
 heat shock protein msbB [imported] - Escherichia coli (strain 0157:H7, substrain R1MD  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C:Accession: E90949  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8: 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and 9  
 A:Reference number: A9629; MUID:21156231; PMID:11258796  
 A:Accession: E90949  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <HAY>  
 A:Cross-References: GB:BA000007; PTDN:BA835988.1; PTD:q13362033; GSPDB:GND0154  
 A:Experimental source: strain 0157:H7, substrain R1MD 0509952  
 C:Genetics:  
 A:Gene: Ecs2565

alignment\_scores:  
 Quality: 127.00 Length: 274  
 Ratio: 0.914 Gaps: 11  
 Percent Similarity: 50.730 Percent Identity: 24.088

alignment\_block:  
 US-09-303-518D-571 x E90949 ..

Align seg 1/1 to: E90949 from: 1 to: 323

```

121 CGCGTGGACATCGCGCTTTTACCTTTAAGGAGACCGCGGCAT 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 Argleucllyargllealagllyargleugllysserserarglrgl 66
171 CGTGGCCATATG.....CGGC 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
66 aleuclleasnlserserleucyspheroargllyargserglualalarg 83
188 AGCGGGTTTGAACCCCGACACGACGTCACGCTTTCGCGGAA 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 lualla.....llevalaspelumelephalalnr 92
238 ACGGCAAAATGCGGTTTGAATTCGCGCGCTTTTCAAAAACCGGA 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
93 Alaproglnalmetalaalmetalaalglulalaleargllyprogl 109
288 AGACATGGAACAATGTTCAAGCGGTACACGCGCTGGGACACGTCGAC 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109 ulyslleglproarglasp...trpglngllyleuglullelleleug 125
338 AGCGTTTGACAGGCGGAGGCGGTGCTGTCACGCGGACATCGCG 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 lmetarglrglansnglulysvalillepheleuvailprohlsgltyr 141
388 AGCTACGATTTGGCGGACGCTTACATGACGACGCGCTTCGTCACCT 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 Alavalaspileroalalmetalealaserln...gllylnyswe 157

```

```
438 GACCGCATGTACAGCCGCGAAATCAAGCATAGACAAATCATGC 487
      ::::::::::::::::::::
157 tAlaAlaMetPheNHisAsnGInGlyAsnProValPheAspTyrValTPrA 174
      ::::::::::::::::::::
488 AGCGGGGAGGGGTGGCGGCAAGGCAAAACCGCCCGGCGCATACAA 537
      ::::::::::::::::::::
174 sTThValAlaTgArGArGpHeClYgLaTgLeuNHisAlaTgAsn...Asp 189
      ::::::::::::::::::::
538 GGGGTCAAAATCATCAAGCGCCCTGGCGGCGGCGGAGCAACATCAT 587
      ::::::::::::::::::::
190 GlyIleLysProPheIleGInSerValAlaTgInleIYtTrrpGlyTyrTy 206
      ::::::::::::::::::::
588 CCTGCC.....GACCACTGCTCTTCTCCGAGAAAGCGCGCGCGTGT 631
      ::::::::::::::::::::
206 rLeuProAspGlnAspHisGlyProGlnHisSerGlu.....P 219
      ::::::::::::::::::::
632 GGGCGGATTTTTCGGCAACCTGCATACACATGACATCGCGGCAAAA 681
      ::::::::::::::::::::
219 heValAspPhePheAlaThTyrIYsAlaThrLeuProAlaIleGlyArg 235
      ::::::::::::::::::::
682 TTGGCACACGTC.....AAAGCGGTGMAAACCTGTTTTCCTGTCGGA 725
      ::::::::::::::::::::
236 LeuMetLysValLysAspAlaArgValValProLeuPhe..... 248
      ::::::::::::::::::::
726 ACGGCTGGCGGAGCGGCAAA.....GGCTTGCTGTGTCACATCGCGCG 769
      ::::::::::::::::::::
249 ProIleLysTgAspLysTgThrNHisArgLeuThrIleGlnValArgPro 265
      ::::::::::::::::::::
770 TCCAGGCGGAATTCAGCGCAACAAACCCACAT...GCCGCGGTTC 816
      ::::::::::::::::::::
265 roMetAspAspLeuLeuGlnAlaAspAspHisThrIleAlaArgGmet 281
      ::::::::::::::::::::
817 AACCGCAATACCGAATATTCGATACGCGCGTTCGAGCGCATATCTGT 866
      ::::::::::::::::::::
282 AsnGlnGlnValGlnIlePheValGlyProArgProGlnIleTyrThr 298
      ::::::::::::::::::::
867 TATGTACAAACCGCATTAACG 888
      ::::::::::::::::::::
298 rIleLeuLysLeuLeuLysThr 305
      ::::::::::::::::::::

seq_name: pIrr2:A85798

seq_documentation_block:
suppressor of htrb, heat shock protein [imported] - Escherichia coli (strain 0157:H7, su
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: A85798
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85798
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-323 <STO>
A:Cross-references: GB:AE005174; NID:912515911; PIDN:AAG56845.1; GSPDB:GN00145; UWGP:Z29
A:Experimental source: strain 0157:H7, substrain EDL933
C:Genetics:
A:Gene: msbB

alignment_scores:
Quality: 127.00 Length: 274
Ratio: 0.914 Gaps: 11
Percent Similarity: 50.730 Percent Identity: 24.088

alignment_block:
US-09-303-518D-571 x A85798 ..

Align seg 1/1 to: A85798 from: 1 to: 323
```

```
121 CGGCTCGACATGTGGCGTTTACCTTTTAAAGACCGCGGCGCAT 170
      ::::::::::::::::::::
50 ArgLeuLysValTgIleAlaTgArgLeuGlyLysSerSerTgArgTgArg 66
      ::::::::::::::::::::
171 CGTGGCAATATG.....CGGC 187
      ::::::::::::::::::::
66 aleuIleAsnLeuSerLeuCysPheProGlnArgSerGlnAlaGlnArg 83
      ::::::::::::::::::::
188 AGCGGGGTTTGAACCCGACAGCGAGAGCGTCAAGCCGTTTTCGGGAA 237
      ::::::::::::::::::::
83 lAla.....IleValAspLleMetPheAlaThr 92
      ::::::::::::::::::::
238 AGCGCAAAATGCGGTTTGGACTTGGCCCGGCTTTTAAAGACCGGA 287
      ::::::::::::::::::::
93 AlaProGlnAlaMetAlaMetMetAlaGlnLeuAlaIleArgGlyProG 109
      ::::::::::::::::::::
288 AGACATCGAAACATGTTCMAAGCGGTACAGCGCTGGAGAACGTTGAGC 337
      ::::::::::::::::::::
109 ulysIleGlnProAlaValAsp...TrrGlnIleuGlnIleIleGlnG 125
      ::::::::::::::::::::
338 AGCTTTGGACAGAGCGGAGGCGGTGCTGTCATCAGCGCGCATCGGC 387
      ::::::::::::::::::::
125 lMetArgTgAsnAsnGlnLysValIlePheLeuValProHisGlyTrr 141
      ::::::::::::::::::::
388 AGTACGATTTGGCGGAGCGGTACATCAGCGAGCGTTCGCTTCCACT 437
      ::::::::::::::::::::
142 AlaValAspIleProAlaMetLeuMetAlaSerGln...GlyGlnLys 157
      ::::::::::::::::::::
438 GACCGCATGTACAGCCGCGAAATCAAGCATAGACAAATCATGC 487
      ::::::::::::::::::::
157 tAlaAlaMetPheNHisAsnGInGlyAsnProValPheAspTyrValTPrA 174
      ::::::::::::::::::::
488 AGCGGGGAGGGGTGGCGGCAAGGCAAAACCGCCCGGCGCATACAA 537
      ::::::::::::::::::::
174 sTThValAlaTgArGArGpHeClYgLaTgLeuNHisAlaTgAsn...Asp 189
      ::::::::::::::::::::
538 GGGGTCAAAATCATCAAGCGCCCTGGCGGCGGCGGAGCAACATCAT 587
      ::::::::::::::::::::
190 GlyIleLysProPheIleGInSerValAlaTgInleIYtTrrpGlyTyrTy 206
      ::::::::::::::::::::
588 CCTGCC.....GACCACTGCTCTTCTCCGAGAAAGCGCGCGCGTGT 631
      ::::::::::::::::::::
206 rLeuProAspGlnAspHisGlyProGlnHisSerGlu.....P 219
      ::::::::::::::::::::
632 GGGCGGATTTTTCGGCAACCTGCATACACATGACATCGCGGCAAAA 681
      ::::::::::::::::::::
219 heValAspPhePheAlaThTyrIYsAlaThrLeuProAlaIleGlyArg 235
      ::::::::::::::::::::
682 TTGGCACACGTC.....AAAGCGGTGMAAACCTGTTTTCCTGTCGGA 725
      ::::::::::::::::::::
236 LeuMetLysValLysAspAlaArgValValProLeuPhe..... 248
      ::::::::::::::::::::
726 ACGGCTGGCGGAGCGGCAAA.....GGCTTGCTGTGTCACATCGCGCG 769
      ::::::::::::::::::::
249 ProIleLysTgAspLysTgThrNHisArgLeuThrIleGlnValArgPro 265
      ::::::::::::::::::::
770 TCCAGGCGGAATTCAGCGCAACAAACCCACAT...GCCGCGGTTC 816
      ::::::::::::::::::::
265 roMetAspAspLeuLeuGlnAlaAspAspHisThrIleAlaArgGmet 281
      ::::::::::::::::::::
817 AACCGCAATACCGAATATTCGATACGCGCGTTCGAGCGCATATCTGT 866
      ::::::::::::::::::::
282 AsnGlnGlnValGlnIlePheValGlyProArgProGlnIleTyrThr 298
      ::::::::::::::::::::
867 TATGTACAAACCGCATTAACG 888
      ::::::::::::::::::::
298 rIleLeuLysLeuLeuLysThr 305
      ::::::::::::::::::::

seq_name: pIrr2:T43023

seq_documentation_block:
ceoA protein - Burkholderia cepacia
C:Species: Burkholderia cepacia
```





```

334 CAGCAGCGCTTTGACAGAGGCGGAGGCTGCTGTTTCATCAGCGGCACAT 383
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
142 rlysaargalathrthrargargasn.....ProThrArgArgThrPT 156
384 CG.....GCAGCTACGATTTGGCGGAGCGGTACATCAGCGCAGC 424
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
156 hrThArgArgAlaProThrLysArgAlaThrThrLysArgAlaThrThr 172
425 TTCGCTTCAGCTGACCGCATGACAAAGCGCGGAAATCAACAGCATA 474
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
173 ArgArgasnProThrLysArgLysThrThrArgArgThrThrValArgAl 189
475 GACAAATATCATGACAGCGGCGGAGGTGCGGCAAAAGCAAAACCGCGC 524
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
189 athrLysThrThrLysArgAla.....ThrThrLysArgA 201
525 CACGCGCTACAGGGGCAACAAATCATCAAGCGCTCGCGCGGCGG 574
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 lAprOThrLysArgAlaThrThrLysArgAla...ProThrLysArgVal 216
575 .....AGCAACCATCATCTGCGCCAGCAGCGTCCCTTCGCGAG 615
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
217 ThrThrLysArgAlaProThrLysArgAlaThrThrLysArgAlaProth 233
616 GAAGCGCGCGCGCGTGGCGGATTTTTCGCGCAACTGCATACACAT 665
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233 rLysArgAlaThrThrLysArgAla..... 241
666 GACACTGCGCGCAAAATGCGACAGCGCAAGCGGTGAACCCGTGTTT 715
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
242 .....ProThrLysArgAla..... 246
716 TCTGCTCGACAGCGCTGCGCGAGCAAGCGTCTGTTGCAATCCGC 765
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
247 ...ThrThrLysArgAlaProThrLysArgAlaThr...ThrLysArgAl 261
766 CCGCTCCAGCGGCAATGAAG.....GCACAAAGCGCCAGCGATGC 806
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
261 aProThrLysArgAlaThrThrLysArgAlaThrAlaArgProThrSerL 278
807 CGCGCGTGT 814
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
278 yspProcys 280
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
seq_name: p1r2:A42608

seq_documentation_block:
(kdo)2-(lauroyl)-lipid IV4 acyltransferase (EC 2.3.1.-) - Escherichia coli
N:Alternate names: membrane-bound lytic transglycosylase
C:Species: Escherichia coli
C:Date: 04-Mar-1993 #sequence revision 18-Nov-1994 #text-change 21-Jul-2000
C:Accession: A42608; 154979; 664947
R:Karow, M.; Georgopoulos, C.
J. Bacteriol. 174, 702-710, 1992
A:Title: Isolation and characterization of the Escherichia coli msbB gene, a multicopy s
A:Reference number: A42608; MUID:92121107
A:Accession: A42608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <R>
A:Cross-references: GB:M87660; NID:q146876; PIDN:AAA96706.1; PID:q146877
A:Experimental source: strain K-12, substrain W3110
R:Engel, H.; Smink, A.J.; Van Wijingaarden, L.; Keek, W.
J. Bacteriol. 174, 6394-6403, 1992
A:Title: Morein-metabolizing enzymes from Escherichia coli: existence of a second lytic
A:Reference number: 154979; MUID:93015688
A:Accession: 154979
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-323 <R>
A:Cross-references: GB:M87660; NID:q146876; PIDN:AAA96706.1; PID:q146877
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cd

```

```

A: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: G64947
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-323 <R>
A:Cross-references: GB:AE000279; GB:U00096; NID:q1788154; PIDN:AACT4925.1; PID:q17881
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: msbB; mlt
A:Function:
A:Description: transfers myristate or laurate, activated on ACP, to (kdo)2-(lauroyl)-
A:Pathway: lipopolysaccharide core biosynthesis
A:Note: decanoyl, palmitoyl, palmitoleyl and (R)-3-hydroxymyristoyl-ACP are poor acy
C:Keywords: acyltransferase; lipopolysaccharide core biosynthesis; transmembrane prot
F:23-43/Domain: transmembrane #status predicted <TM2>
F:85-105/Domain: transmembrane #status predicted <TM2>
F:133-153/Domain: transmembrane #status predicted <TM3>

```

```

alignment_scores:
      Quality: 125.00      Length: 274
      Ratio: 0.899      Gaps: 11
      Percent Similarity: 50.730      Percent Identity: 24.088

```

```

alignment_block:
US-09-303-518D-571 x A42608

```

```

Align seg 1/1 to: A42608 from: 1 to: 323

```

```

121 CGCGCTGACATCTGCGGTTTACCTTTTAAGAGACCGCGCGCAT 170
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
50 ArgLeuGlyArgPheAlaGlyArgLeuGlyLysSerSerArgArgAl 66
171 CGTGGCAATATG.....CGCG 187
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
66 alauIleasnLeuSerLeuCysPheProGluArgSerGluAlaGluArgG 83
188 AGCGGGGTTTGAACCCGACAGCGAGGTCGAAGCCGTTTTCGGGA 237
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
83 luAla.....IleValAspGluMetPheAlaThr 92
238 ACGGCAAAATGCGGTTTGAATTCGCCCGCGGTTTTCAAAACCGGA 287
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
93 AlaProGluAlaMetAlaMetAlaGluLeuAlaIleArgGlyProG 109
288 AGACATCGAAGACATGTTCAAGCGGTACAGCGGTGGGAGACGTCGAGC 337
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
109 uLysIleGlnProAlaArgValAsp...TyrPheGlyLeuGluIleIleG 125
338 AGGCTTTGGACAAGGCGAAGGGGTGCTGTTTCATCAGCCGACATCGGC 387
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
125 lmetArgArgAsnAsnGlnLysValIlePheLeuValProHisGlyTyr 141
388 AGCTACGATTTGGCGGAGCGGTACATCAGCGAGAGCTCCGTTCCACT 437
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
142 AlaValAspIleProAlaMetLeuMetAlaSerGln...GlyLysLys 157
438 GACCGCCATGTACAAAGCGCGGAAATCAAGCGATGACAAATCATGCG 487
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
157 tAlaIleMetPheHisAsnGlnLysProValPheAspTyrValTyrP 174
488 AGCGCGGAGGGTGGCGGCAAGGCAAAACCGGCCACCGCATACAA 537
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
174 smThrValArgArgPheGlyArgGluHisAlaArgAsn...Asp 189
538 GGGGCAAAATCATCAAGCGCGTGGCGGCGGAGGAGCAACATCAT 587
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
190 GlyIleLysProPheIleGlnSerValArgGlnGlyTyrTyrPlyTyr 206
588 CCGCGCC.....GACCAAGTCCCTTCTCGCGAGGAAGCGGCGCGTGT 631

```

